A Clonal Selection Algorithm and Extensions

JASON BROWNLEE
Technical Report 070521A
Complex Intelligent Systems Laboratory, Centre for Information Technology Research,
Faculty of Information and Communication Technologies, Swinburne University of Technology
Melbourne, Australia
jbrownlee@ict.swin.edu.au

***Abstract*-The clonal selection theory suggests an information processing principle of the selection and proliferation of temporally useful information. The proliferation strategy promotes maturation of response via blind mutation and anticipation of selection through adjusted information densities. This work explores the clonal selection principle as an information-processing algorithm and proposes (1) a minimal approach, (2) an elaborated approach, and (3) implications and extensions of the proposed system.**

Keywords-Clonal Selection Theory, Algorithm, Artificial Immune System

I. INTRODUCTION

The clonal selection principle is an abstraction of the clonal selection theory an describes a stimulus response adaptive plan in which antigen select cell receptors based on affinity resulting in clonal expansion, maturation by hypermutation and cell differentiation. This work proposes a simple clonal selection algorithm based on this adaptive plan. Natural extensions to this simple algorithm are enumerated and fundamental concerns with the approach are addressed such that the result is a class of information-processing algorithms.

Section II summarises some design principles in devising a clonal selection algorithm. Section III proposes a minimal clonal selection algorithm that considers the clonal selection principle and the outlined design principles. Section IV elaborates on the minimal algorithm by expanding some of the algorithm steps into sub-tasks and in adding an additional algorithm step. Section V discusses the implication of emergent cells classes in the elaborated algorithm, and proposes a number of natural extensions of the approach.

II. SOME DESIGN PRINCIPLES

This section provides an ensemble of design principles that influence the proposal and extension of the clonal selection algorithm.

The clonal selection principle is (1) the external identification of useful patterns, (2) the internal duplication, and modification of identified internal patterns. The system is concerned with providing the best possible response to an antigenic query and internally responding to the query through a process of blind adaptation.

Principle: *External stimulation which results in internal activation that takes the form of blind maturation of response*

The immune system is envisaged as being situated in an environment where the task to perform is pattern recognition. This pattern recognition task is achieved using an oracle function to which system patterns are submitted, and amorphous affinity scorings returned. In this regard, affinity is an extrinsic property of the system; it is assigned externally, and implies the temporal usefulness of a pattern for an unknown reason. The application of the oracle function may be conceptualised as the pointing out of some system patterns that are currently meaningful, to which the system must respond.

Principle: *Affinity is inherently extrinsic and only has meaning for a given antigen, and a give receptor, at a given time*

The system is allocated limited resources in which to store the repertoire of receptors, thus the resources allocated (space complexity) must be used effectively in the context of the problem. Further, the direct evaluation of patterns is also a moderated resource that must be used sparingly (time complexity), such that the system is capable of adaptation in the absence of additional external pattern evaluation.

Principle: *Space and time complexity are limited resources (cost more than the internal operations of the system), thus must be used efficiently to promote the efficacy of the system in the domain*

A binary encoding for pattern representation provides a flexible basis for investigating the clonal selection principle. A bit string facilitates genetic-like operators such as copying errors during duplication and transformation via transcription for mapping into domains. In addition, pattern recognition properties are easily realised such as genetic similarity (binary distance), binding (matching function), and generation of naïve patterns (random bit strings).

Principle: *Use of a binary representation for pattern recognition*

An effective realisation of the clonal selection principle is a system in which relevant cognitive properties of pattern recognition are self-regulating in response to the regularities of the domain. Concerns such as memory, acceptance of novelty, clonal dominance, receptor densities, and maturation of response are emergent properties of an effective realisation.

Principle: *A self-regulating system with emergent cognitive properties*

Some adaptive tools available to the clonal selection principle include competition between receptors for recognition by antigen, attrition in the death and removal of unused receptors, and the proliferation and blind maturation of useful receptors. Information available to these tools include temporal usefulness of receptors (both current and past) and the genetic similarity of receptors.

Principle: *Tools include competition, attrition, and maturation in proliferation, available information includes temporal usefulness and genetic similarity of receptors*

The effective use of these tools and information will result in pressures in the system that the guide behaviour. The effective tuning and application of these pressures define the behaviour (efficacy) of the system in a domain as well as the performance of the system (space and time complexity). Perhaps the most important pressure in the system is competition, which includes competition for clonal selection, competition for survival, and competition for replacement into the repertoire.

Principle*: Competition is the guiding pressure of the system*

The system is anticipatory in principle. The outcome of the proliferation and maturation is the assumption that patterns that are currently and temporally useful, will be useful in the future. Thus, the economy of this anticipation is payoff, the exploitation of patterns that are most likely to payoff (be useful) in the systems future exposures to antigen.

Principle*: The system anticipates useful patterns guided by the pursuit of expected payoff*

A given receptor is redundant, such that it may be lost without major detriment to the systems performance. In addition to this information redundancy, the cells do not live forever; rather they are continually turned over. Unused cells are removed in a form of attrition, active cells are employed in the response, and there is a constant stream of new cells created as a function of exposure. Naïve cells are also continually introduced into the system.

Principle*: A given receptor is redundant and the repertoire is in state of constant flux*

The clonal selection principles describe a combined anticipatory learning scheme that is coupled with a defensive response mechanism. Learning, adaptation, and memory formation occur online (as opposed to offline) at the same time as the system must neutralise pathogen. Thus, the scheme embodies the efficiency and efficacy tradeoffs associated with online learning.

Principle*: The clonal selection scheme describes an online learning and response system*

III. A Minimal Clonal Selection Algorithm

This first algorithm represents an exercise in reduction, specifically the reduction of the clonal selection principle to the simplest possible instantiation as an algorithm. The central theme of this algorithm is iterative selection and mutation upon a population. The algorithm is important because it demonstrates some fundamental induction characterises of the principle, and starkly highlights some problems with such an minimalist realisation.

Step1: *Initialise* the repertoire with a set of randomly generated bit strings
Step2: *Expose* the system to a randomly selected antigen
Step3: *Select* the highest affinity bit string from the evaluated repertoire
Step4: *Clone* the best matching cell with genetic copying errors
Step5: *Replace* the best matching cell with the clone
Step6: *Repeat* steps 2-5 until a stopping condition is triggered

Figure 1 - A minimal clonal selection algorithm

The algorithm presumes the existence of a set of antigen patterns and a repertoire of receptor patterns. It also presumes an affinity mapping function for an antigen and receptor, and genetic operators for cloning of a receptor.

Antigenic Set: A set of antigen patterns that are exposed to the system. May be static or dynamic, known or unknown, binary or some other representation. The order in which they are exposed to the system is presumed random, although may be any distribution over the antigenic set. Further, the algorithm only performs a cycle on the availability of antigen.
Antigen: A single pattern to which the system is exposed. This external pattern is not directly observable by the system, rather it is interacted with indirectly by the receptors in the repertoire via the affinity mapping function.
Repertoire: Set of internal system patterns that encapsulate the information induced by the system processes. Receptors do not interact directly with the environment, thus input patterns never leave the system. They are created, used, and destroyed in the repertoire.
Receptor: A single pattern in the system. Receptor patterns interact with antigen patterns indirectly via an affinity mapping function, although they may interact with each other directly based on intrinsically available heuristics.
Affinity Mapping Function: The Antigenic set and the repertoire are decoupled and interact indirectly via the affinity mapping function which acts as mediator. The mapping function may or may not transform receptor patterns and antigen patterns such that an affinity may be assigned. The affinity assigned has meaning for a given antigen, and a given receptor at a given time.
Genetic Operators: The genetic operators are used in the cloning of high affinity receptors. They are responsible for the duplication of genetic material of a given receptor with copying errors. These errors may include insertions, deletions, inversions, and point mutations.

Figure 2 - Summary of the principle characteristics of the minimal clonal selection algorithm

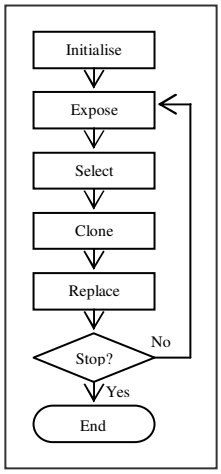

Figure 3 – Flow chart of the minimal clonal selection algorithm

The repertoire of receptors will rapidly stabilise in response to exposed antigen. The winner-take-all selection of receptors by antigen will result in the specialisation of receptors in the repertoire for antigen. The default acceptance of clones may result in the receptors not becoming too specific to the antigen (maintaining some generality), further the acceptance regardless of improvement will influence the systems speed of convergence with a chronic antigen.

Best Response: The winner-take-all (WTA) principle employed in responding to antigen ensures that the system provides its best possible response for a given exposure.
Stable: The WTA principle will likely cause the system to stabilise with regard to consistency of response. This stabilisation process will be rapid, reinforced by the specialisation of receptors for antigen.
Specialisation: The repertoire will self-organize such that receptors specialise for a given antigen. This specialisation will be reinforced resulting in the stabilisation of the repertoire and the continuous iterative adjustment of receptors for their associated receptor.
Moderated Convergence: The acceptance of clones irrespective of their affinity to the antigen as replacements for the highest matching affinity will result in a convergence that is moderated by random jumps. The system will converge (stabilise), but will do so in the context of a continual local search around best matching receptors.

Figure 4 - Summary of some properties of a minimal clonal selection algorithm

A consequence of the pattern-exploitative properties of the winner-take-all principle is that some receptors in the repertoire may never win. The effect of not winning is that a portion of the limited resources employed by the system are unused. For some domains, this may be an efficient side effect, and those receptors that are not used for a time may be removed, leaving a repertoire of useful receptors. Taken to an extreme, a particularly general receptor may always win. The effect is that the receptor (or small set of receptors) will indirectly (directed random) seek an average of the patterns to which they respond. These general receptors will handicap the system, and prevent the specialisation of receptors to differing patterns. The required replacement means that untested clones (with errors) supplant tested patterns (which we know something about).

Concern: *Winner take all will create dominant winners resulting in ineffective application of available resources and handicapped performance*

The implied solution is to raise a set of receptors for each receptor, to address a given antigen from multiple and (likely) different perspectives. This solution is facilitated by the underutilised resources of the winner-take-all principle. These resources may be considered holes in activity (inactivity), and may be plugged with these so-called 'different perspective' receptors. The obvious effect of such a behavioural modification is the slowing of system convergence, and system instability. The intended benefit is the increased utilisation of resources and increased performance. All of these factors may be measured and compared.

Convergence: A measure of time (algorithm cycles, epochs of the antigenic set) until major changes to the repertoire (receptor genetic composition) and system response (affinity to antigen) stabilise
Stability: Amount of variation in repertoire composition (receptor genetic composition) and system response (affinity to antigen) after convergence
Performance: The systems ability to respond to an antigenic set (affinity)
Utilisation: The frequency of use (time in algorithm cycles since last use) of all receptors in the repertoire

Figure 5 - Summary of four base measures in which to compare extensions of the minimal clonal selection algorithm

IV. ELABORATIONS ON THE ALGORITHM

This section proposes four different ways of changing the minimal clonal selection algorithm to address the principle concern of the approach: improve resource utilisation by addressing a given antigen from multiple perspectives.

A. *Elaborate Selection*

The algorithm may be rephrased in between the Step3 (selection), and Step4 (cloning). The selection step implies (at least) three tasks, which may be elaborated:

1. The evaluation of the repertoire of receptors within the context of a given antigen (calculate affinity)
2. The returning of the highest affinity receptor to meet the needs that the antigen represents (a query, a threat, etcetera)
3. The identification of a high-affinity set of receptors that are currently useful and may be made more useful (via cloning)

Thus the minimum clonal selection algorithm may be redefined as follows:

Step3a: *Evaluate* the repertoire in the context of the selected antigen
Step3b: *Return* the highest-affinity antigen to meet the needs of the antigen
Step3c: *Selec*t a high-affinity set of activated receptors

Figure 6 - Modified clonal selection algorithm with an activated set

The change to the algorithm presumes the presence of an activated set, which in the minimal clonal selection algorithm consists of a single receptor: the highest affinity receptor for the exposed antigen.

Activation Set: A collection of receptors that have high affinity for a given antigen at a specific exposure time, which are subjected to cloning.

Figure 7 - Summary of the principle components of the change to the selection set in the minimal clonal selection algorithm

The winner-take-all principle is defined by the highest affinity receptor addressing an antigen, but also being the subject of adaptation (getting it all). In enlarging the activation set, the winner does not get all, but instead, gets most (of the computation). Further, the activation set may or may not possess the highest affinity receptor, thus the winner may only get half of the outcome of the exposure. The increased size activation set changes the algorithm from monoclonal activation to polyclonal clonal activation, allowing the repertoire to address a given antigen from multiple perspectives (runners up or competitors). A further step in addressing the concern with the minimal algorithm, is that the activation schedule is likely to ensure that the algorithm utilises more resources. The following outlines some schemes for selecting an activation set:

Winner-take-all: (WTA), the baseline for comparison as employed in the minimal clonal selection algorithm. The activation set contains only the highest-affinity receptor.
Greedy: The set is selected as a collection of the highest affinity receptors

Probabilistic: The set is selected probabilistically with a bias towards higher affinity receptors (such as with a ranking, tournament or biased roulette wheel method).
Random: The set is selected using a uniformly random distribution function.

Figure 8 - Proposal of some activation set selection schemes

The three proposed selection schemes require a user parameter, which defines the size of the set. Given the cloning step, a given receptor may not enter the set more than once (no reselection). If the best matching receptor is not added to the activation set, it will not be replaced and the information it defines is retained. A further implication is that there is also no chance that that specific piece information (winner) may be improved upon. Activation is a reward – it is the verification that a piece of information (receptor) is presently useful, thus, the activation set must be small (unless everything is useful).

***Oligoclonal Principle**: The number of receptors selected in the active set is small relative to the size of the repertoire.*

Another concern is that the activation set may possess receptors that are not relevant to the antigen. This may be a desirable effect in that receptors that are not used often may be adjusted such that they become useful, and the system may become more responsive (increased performance) for a novel antigenic pattern. Alternatively, a consequence is that information that is desirable for other purposes (such as for another antigen) may be lost at that time.

***Concern**: An alternative selection scheme of the activation set may result in the right information not been maturated, and the wrong information being maturated*

The cloning and maturation process is antigen invariant, so it is possible that the maturation of irrelevant antigen may have a beneficial outcome. The concern from a computational perspective is that computational resources may be employed unnecessarily.

B. ***Elaborate Cloning***

The cloning step (Step4) may be rephrased, and elaborated into the following set of sub-tasks for each member of the activated receptor set:

1. Create a clone of genetically duplicate receptors
2. Impose copying errors and mutations on the clone of receptors
3. Select a clonal set of clones which may replace members of the activated set

Thus, the minimal clonal selection algorithm may be redefined as the following:

Step4a: *Clone* each receptor in the activated set to create a clonal set
Step4b: *Mutate* each receptor in the clonal set for a given activated receptor
Step4c: *Select* a subset of each clonal set to compose a replacement set

Figure 9 – Modified clonal selection algorithm with elaborated cloning

The modified algorithm presumes the construction of a clonal set for each member of activated set, a mutation function, which is applied in the creation of members of the clonal set (copying errors), and the selection of members of each activated receptors clonal set which in aggregation becomes the replacement set.

Clonal Set: A set of cloned receptors with a common genetic ancestor, which is a member of the current activated set. The clone is created as duplicates (clones) of the activated progenitor with copying errors.
Mutation Function: Employed to create genetic duplicates of a given activated receptor, which may introduce copying errors such as additions, deletions, inversions, and point mutations of the binary string
Replacement Set: A set of clones selected from each activated receptors clonal set. This set replaces each of the members of the activated set, thus must contain the same number of receptors.

Figure 10 - Summary of the components of the elaborated cloning procedure

In the case of the minimal algorithm, the activate set contains the highest affinity receptor. The clonal set created for this receptor contains a single clone created with copying errors, which is in turn selected into the replacement set.

A single activated receptor may fill the clonal set with an arbitrary number of clones (parameterised), each (likely) with a different genetic composition given coping errors. The number of clones could be a function of the relative affinity of the progenitor to other receptors facilitating a '*higher the affinity, the more clones created*' principle which may be useful in distinguishing the clonal rate of receptors in an activated set larger than one.

Parameterised: The number of clones created for an activated receptor is defined by an arbitrary parameter.
Affinity Proportional: The number of clones created for an activated receptor is proportional to the affinity of the receptor to other receptors. For an activation set of one WTA), the rate may be a function of the receptors affinity delta to the second highest affinity receptor. For an activation set of >1 or non WTA principle, the clonal rate may be a proportionate function of the receptors affinity rank in the repertoire or activated set.

Figure 11 - Summary of some cloning schemes for defining the size of an activated receptors clonal set

Copying errors occur during the cloning of a receptor the effect of which are changes to the clones genetic composition compared to its progenitor. Although the composition is changed, the clonal selection principle suggests that progeny are within a neighbourhood of the progenitor. Thus, the copying errors made during cloning have a high probability of creating a receptor similar (duplicate) to the progenitor, and a low probability of creating a receptor vastly different from the progenitor. A copying error may not define the amount of conformation difference, but rather the amount of genetic difference: mistakes in replicating genetic material. Making multiple clones increasing the chance of a copying error, but dramatically increases the chance of making a receptor without a copying error.

***Cloning Principle**: The more clones created, the more chance of a copying error, and the much larger chance of creating duplicates of the original receptor.*

A duplicate receptor represents retention of information, the small modification of a receptor represent the gain (improvement) or loss (noise) of existing information, and large modification of a receptor represents the loss of specific information and the potential gain of new information.

Parameterised: The amount of mutation may be defined by an arbitrary parameter (such as a mutation probability). In the case of affinity-proportionate clonal set size, the affinity would also influence the number of total mutations and number of duplicates and close-to-duplicate receptors are created.
Affinity Proportionate: The same relative affinity heuristics may be employed on the amount of copying errors. An inverse relationship may be used such that the higher affinity of a receptor to an antigen, the less likely copying errors are to occur in creating the receptors clonal set.

Figure 12 - Summary of the mutation schemes

The selection of a replacement set is the selection of receptor clones that will replace the activation set. In the case of each activated receptors clonal set holding a single clone, no selection process is required, it is only when the aggregation of clonal sets is larger than the activated set that reduction of the union is required. It is not the task of this selection task to replace the activated set, but rather to propose a clonal set, which may be used in the replacement step. The default action for this selection step, is thus to provide the union of clonal sets as the replacement set.

Select All: The union of the clonal sets is taken as the replacement set, and the size difference between the replacement set and the active set is reconciled (or not) by the replacement steps integration into the replacement set in the repertoire
Similarity: The genetically most similar receptor of each clonal set to its progenitor is selected into the replacement set
Probabilistic: Receptors may be selected probabilistically from each clonal set with a bias towards genetic similarity (or dissimilarity) to each clonal sets activated receptor. Alternatively, receptors may be drawn from clonal sets proportional to the relative affinity of each clonal sets active receptor.
Random: Receptors may be drawn randomly from the union of the clonal sets

Figure 13 - Summary of the receptor set selection schemes

Cloning, mutation, and replacement set selection do not naturally impact on the primary concern of the minimal clonal selection algorithm. In replacing receptors with high approximations (duplications or near duplications) will assist the WTA principle by maintaining a stable repertoire. Replacing the active set with low receptor approximations will dampen the WTA principle and promote the utilisation of the broader repertoire, without any adaptive capability.

The fidelity of active receptor clones does directly influence information retention, and the acquisition and rate of acquisition of new information (learning).

C. *Elaborate Replacement*

The replacement step (Step5) may be rephrased, and elaborated into the following tasks:

1. The selection of a replacement (supplant) set from the repertoire to complement the replacer set (derived from clonal sets)
2. Replace the supplant set with the replacement set

Thus, the minimal clonal selection algorithm may be redefined as follows:

Step5a: *Select* a subset of the repertoire to compose the supplant set
Step5b: *Replace* the supplant set with the replacement set

Figure 14 – Modified clonal selection algorithm with elaborated replacement

The modified algorithm presumes the construction of a supplant set.

Supplant Set: A set of receptors from the repertoire, which are replaced by the replacement set which is composed of clones of activated receptors. Receptors the supplant set are kill, deleted from the system.

Figure 15 - Summary of the principle components of the elaborated replacement mechanism

The default supplant set is the activated set of receptors, the progenitors of the clones to which they are being supplanted. The elaboration of the replacement step allows the selection of receptors from the repertoire to be replaced by clones of the activated set.

Activate Set: The active set (progenitors) are selected as the supplant set
Random: Random receptors may be drawn from the repertoire such that the size of the supplant set matches the replacement set.

Figure 16 - Supplant set selection schemes

The capability to choose to not replace progenitor receptors, the repertoire is provided with the ability to both retain demonstrated high-affinity receptors, and potentially higher-affinity clones of the receptor. In replacing repertoire members that have not been activated for a time (underutilised), this amendment to replacement directly addresses the concern of the minimal clonal selection algorithm. A consequence of this behaviour, is that receptors for other antigen (that are useful) may also be replaced.

Concern: *In clones not replacing progenitors it is possible for information presently unrelated (to the current antigen) to be replaced*

The supplant set defines the members of the repertoire that the replacement set compete with for a position in the repertoire. The default form of competition is no competition – the replacement set always wins against the activation set. Alternatively, with a competition perspective, a potential supplant set may be selected for each member of the replacement set.

Similarity: A replacement receptor may compete with a supplant set based on genetic similarity. The clones progenitor is likely to be the genetically most similar receptor in the repertoire, and if the supplant set is drawn per-replacement receptor, clonal siblings (if there are any) are also likely to be the most similar receptors in the repertoire.
Affinity: A replacement receptor may assume the affinity of its parent activated receptor, and thus may compete based on affinity. Affinity similarity may be used to select receptors in the receptor set that that have a similar present utility, although potentially different genetic basis. Affinity may be used as a discriminator in which clones are rejected after the application of another heuristic such as genetic similarity (assuming something other than WTA is used to prepare the active set)
Utilisation: Replacement receptors may compete with members of the supplant set based upon the utilisation of receptors in that set. A default policy of replace poorly utilised receptors may be used.

Figure 17 - Summary of competition schemes for the replacement and supplant set of receptors

Ideally, the replacement scheme will be biased toward replacing (1) genetically similar (2) underutilised receptors with (3) lower affinity than the replacer receptor's progenitor. Each of the three concerns support each other, thus, this may be defined as a multiple objective problem.

D. *Add Naïve Receptors*

A step may be added into the minimal clonal selection algorithm after replacement (Step5) that inserts

naïve (random) receptors into the repertoire. This step involves the following tasks as follows:

1. Generation a naïve receptor set
2. Selection of a supplant set
3. Replacement of supplant set by the replacement set

Thus, the minimal clonal selection algorithm may be redefined as follows:

Step6a: *Generate* a naive set of randomly generated bit strings
Step6b: *Select* a subset of the repertoire to compose the supplant set
Step6c: *Replace* the supplant set with the naive set

Figure 18 – Modified clonal selection algorithm with added naïve replacement

The amendment to the algorithm presumes the generation of a naïve receptor set, the selection of supplant set and competition for replacement.

Naïve Set: A set of randomly generated receptors to insert into the repertoire
Supplant Set: A set of receptors selected form the repertoire to replace with the naïve receptor set

Figure 19 - Summary of the principle components of the naive receptor amendment

The preparation and replacement of a supplant set has many of the same concerns to that of the selection and replacement of a supplant set to be replaced by the replacement set after cloning. Two important differences include the lack of parental affinity associated with the generated naïve receptors, and the knowledge that receptors in the repertoire are naïve. The lack of parent affinity limits the heuristics that may be used in the preparation of the supplant set and in the competition for replacement between the supplant set and the naïve set. The knowledge that a receptor is naïve allows additional features such as the replacement of older unused naïve cell by newer naïve cells. This step may address the concerns of the minimal clonal selection algorithm by replacing those receptors that are underutilised with naïve receptors. This strategy may permit new receptor lines to be investigated for a given antigen, if the naïve receptor lines are able to 'take root'. A concern with this approach is that the utility of naïve receptors may be less than that of the receptors that they are replacing, and the likelihood of a naïve receptor having a higher affinity than a matured affinity is low.

This mechanism may be suited to replace those underutilised receptors that 'fall through the cracks' of the replacement scheme employed by the clonal set based replacement set. In the scenario in which the replacement set performs replacement based on genetic similarity, it is likely that the repertoire may reach a stable point of genetic-similarity-based replacement that leaves genetically dissimilar receptors untouched in the repertoire. In this case, the benefits of a genetic-invariant (utility based) naïve receptor replacement scheme are apparent.

E. *Elaborated Clonal Selection Algorithm*

Given all the elaborations of this section, we may now define a rephrased minimal clonal selection algorithm as an elaborated clonal selection algorithm.

Step1: *Initialise* the repertoire with a set of randomly generated bit strings
Step2: *Expose* the system to a randomly select an antigen
Step3a: *Evaluate* the repertoire in the context of the selected antigen
Step3b: *Return* the highest-affinity antigen to meet the needs of the antigen
Step3c: *Select* a high-affinity set of activated receptors
Step4a: *Clone* each receptor in the activated set to create a clonal set
Step4b: *Mutate* each receptor in the clonal set for a given activated receptor
Step4c: *Select* a subset of each clonal set to compose a replacement set
Step5a: *Select* a subset of the repertoire to compose the supplant set
Step5b: *Replace* the supplant set with the replacement set
Step6a: *Generate* a naive set of randomly generated bit strings
Step6b: *Select* a subset of the repertoire to compose the supplant set
Step6c: *Replace* the supplant set with the naive set
Step7: *Repeat* steps 2-6 until a stopping condition is triggered

Figure 20 - Elaborated Clonal Selection Algorithm

The elaborated clonal selection contains the same primary tasks of the minimal clonal selection algorithm (initialisation, exposure, selection, cloning, and replacement), with specific tasks of each step expanded upon. The elaborated algorithm also has an added naïve cell replacement step.

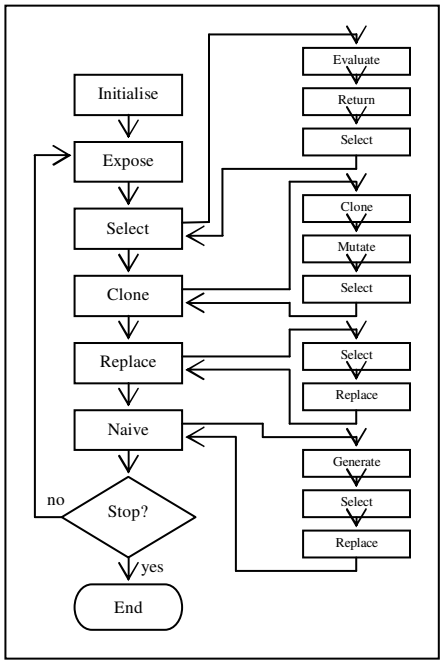

Figure 21 - Elaborated Clonal Selection Algorithm

V. SOME IMPLICATIONS AND EXTENSIONS

This section discusses some implications and extensions to the elaborated clonal selection model and elaborations proposed above. In particular, this section is concerned with the broader concerns when the system is considered holistically with all elaborations enacted.

A. *Emergent Cell Types (Classes)*

It is proposed that the system can evolve emergent cell types through the effective application of selection and replacement heuristics. These cell types are not delineated by anything other than level of maturation for their antigen (affinity) and their usefulness through time (activation count). The result is a continuum of receptors with the following (arbitrarily) defined classes.

Memory Cell: Relatively high affinity for a given antigen, with many

activations in its lifetime, not likely to be replaced by anything other than a competing effector cell on its ascension to memory cell
Effector Cell: Moderate to high affinity receptor for a given antigen with few activations, likely to be replaced by other competing effectors
Naive Cell: A low to moderate affinity receptor for most antigen with none or very few activations, likely to be replaced by an effector cell or another naive cell

Figure 22 - Propose cell types delineated by amount of activation and affinity for their antigen

The desire is for the system to be configured such that competition is between receptors of the same cell type. Memory cells should be useful when competing for their antigen (step3). Effector cells should also be useful in competing for their antigen, although should be the fodder that is replaced by other effector cells for the same antigen, including clones of memory cells. Naïve cells (as well as low affinity effector cells) should be useful in step3 in addressing novel antigen, and provide fodder for low affinity clones, and randomly generated receptors.

***Competition Principle**: Competition at the various levels should be biased to be between information of the same context, the same level of usefulness, and the same age. Alternatively, cell lines created and matured for differing antigen should cooperate (at the system level) rather than compete.*

Memory cells for another antigen (low affinity) still look like memory cells given a history of activations, thus discriminative pressure for memory is durability and usefulness. Effector cells for another antigen (low affinity) are difficult to distinguish between naïve cells. This difficulty implies that clones (new potential effectors) should only compete with other effectors for the same antigen (genetically similar, affinity similar receptors). Thus the supplant selection and replacement scheme (Step5a/b) should be biased toward similar genetic composition, lower affinity, and low activation count. The pressure for effector-effector competition, although biased, will not preclude the replacement of an effector for a different antigen, or a naïve cell, or even a memory cell for the same antigen that has not been used for some time. Further, the stabilising effects of winner-take-all (and variations) in conjunction with the genetic clustering will result in the formation of genetically disparate memory cell lines.

Memory is perhaps the most critical emergent property, facilitated by the ongoing self-organization of the repertoire into the other two types of memory (short term and long term). The emergent memory must be robust enough such that if information that should become memory is discovered, that it very likely for it to become memory. Further, if memory is lost (damage), or becomes irrelevant (pathogen is exterminated) the mechanisms are such that this self-organized repertoire cell hierarchy will adapt.

***Memory Principle**: Cells are memory cells because they are consistently (multiple activations) demonstrated to be useful (high affinity)*

Effectors are variations on relatively and presently high affinity receptors that are proposed to address an immediate concern. If they are not useful to the present antigen (low affinity), assuming the present antigen is chronic, they will be supplanted. If effectors are useful to the present antigen (high affinity), they take residence in the repertoire. If their usefulness persists (multiple activations), they ascend to memory status (emergent).

***Memory Types**: Memory cells (long lifetime of useful activation) represent long-term memory that is difficult to dislodge. Effector cells (short lifetime of moderate to low useful activation) represent short and medium term memory that is relatively easy to dislodge.*

This has interesting implications in that the repertoire and replacement scheme must facilitate the retention of effectors for long enough to allow them to 'take root' if an antigen exposure is moderately infrequent. A natural inclination of the system will be for effector cells to be supplanted by effectors in the context of other antigen (if there are other antigen in the intervening interval), or to be replaced by naïve cells (in the absence of antigen in the interval). The pressures must be finely balanced. The replacement pressures must be such that there is a graceful fall-off in the survivability of useful effectors in the presence of infrequent antigen exposures.

***Effector Principle**: Effectors are effectors because they are created as needed for an antigen, and likely are supplanted by other effectors for the antigen, and after some time, effectors for other antigen, and naïve cells (graceful falloff in that order).*

A chronic infection will result (eventually) in a biased repertoire for that infection (takeover). Antigen may not last forever, and novel antigen may arrive. There are many reasons for naïve cells, although the selection of a suitable time for their arrival may be difficult to determine *a priori*. Naïve cells are those receptors from an untested cell line, thus they themselves are untested, and most likely useless in normal circumstances. In this regard, they are similar to large mutations of high affinity receptors (likely useless), underutilised cells that have not seen activation in a long time (temporally useless), and receptors without activation in their lifetime such as unused effectors and other naïve cells (temporally useless).

Thus, cell types may be delineated by their lifespan in the repertoire. Memory cells remain through long-term usefulness, effector cells linger briefly to address temporal needs, and naïve cells fill in the gaps with a relatively short lifespan.

***Lifespan Principle**: Cell types delineate relative receptor lifespan in the repertoire*

B. *Extensions*

The elaborated clonal selection algorithm is a pattern recognition system in which input patterns are not only recognized, but their existence recorded and organized within the repertoire of receptors. This section briefly describes some natural extensions to the proposed algorithm.

Automatic Configuration

It is proposed that through effective configuration many of the aspects of the algorithm that lend themselves to be parameterised may be addressed

automatically by the processes of the algorithm. Examples include (1) the size of the activated set, (2) the number of clones created and resultant modification of knowledge, and (3) the amount of naïve cells inserted. The first two examples may be addressed through the effective use of genetic composition and affinity deltas as heuristics (a diverse and high affinity response). The third example may be addressed through effective use of activation history and last activation time as heurists for replacement (unused, non-memory).

***Automatic Heuristics**: Parameterisation assumes the meaningful tuning of behaviour, which implies underlying heuristics that govern that tuning. These heuristics may be applied automatically, and dynamically.*

The intent of such automatic configuration is that the system may allocate resources proportional to their assessed requirement. Thus, hyper-activation and or hyper-cloning may occur to address novel antigen or other reasons for poor affinity of response. Clonal dominance (takeover of the repertoire) may occur in a mono-antigenic environment, and hyper-naivety (many naïve receptor replacements) may occur in high noise antigenic environments. A final seminal example is the automatic resizing of the repertoire in response to the diversity of antigenic signals the system is exposed. This would allow the repertoire to expand (bloat) in times of diverse exposures, and to contract via mass-attrition in times of mono-antigenicity.

Elaborated Genetics

Variation is introduced via copying errors in receptor duplication via insertions, deletions, inversions, and point mutations. This mechanism may be elaborated through the use of a common genetic basis from which receptors are drawn. The genetic basis used by each receptor would still be subjected to variation, although the genetic basis is remapped by the receptor. This remapping decouples the affinity scoring from the genetic material and assigns it to the receptors mapping of portions of the genetic material in the form of its conformation. A simple example[1] of this remapping is a simple masking scheme as follows:

Receptor:	[01111]
Conformation:	[_01_1_1_1_]
Mapping:	[#__#_#_#_#]
Genetic:	[1**0**1010101**0**10]

Figure 23 - Example of a more elaborate genetic representation scheme

The base genetic material is mutated by whatever means. The mapping may represent the 'genes' (collections of alleles) in the genome or perhaps a string of amino acids for the protein. The conformation may represent the string of amino acids, or the folded protein, and finally the receptor represents the part of the protein that does the pattern recognition. This scheme facilitates variable length receptors, and mutation of the mapping function, which inturn may also be defined in the genetic material. Variable receptor lengths provides and promotes a mechanism of polyclonal recognition, in which a given receptor may match onto many different antigen that exhibit the same general pattern (cross-reactivity).

***Decoupled Representation**: A more elaborate receptor genetic representation facilitates variable receptor size, recessive receptor information, and higher-order (macro) changes resultant receptors*

A decoupled genetic representation and pattern recognition allows mutations to occur without effect to the receptor (conformation), and receptors to change without changes to the genetic material (mapping). It also facilitates the storage of receptor information without use (in the unmapped genetic representation), and high-order changes to receptors during clonal expansion.

Cell-Cell Recognition

Receptors are pattern recognition devices, and a repertoire of them represents a pattern recognition machine. Receptors themselves are patterns of bit strings, thus a natural extension to explore is the recognition of receptors by other receptors (a concept that is reminiscent of the network theory of immunology). A natural interpretation of cell-cell recognition and interaction is the genetic-based replacement heuristics involved in consolidating genetically similar groups of receptors, providing a regulation mechanism.

***Antigen-Independent Information Processing**: Pattern recognition facilitates information longevity (via increased activation counts and replication), thus antigen-independent processing via intra-repertoire pattern recognition allows the repertoire to reinforce previously acquired information.*

Cell-cell recognition may be used for information processing in the absence of antigen, where the activated receptors themselves are promoted to antigen-like status in the repertoire. Interaction based on genetic similarity would not be suitable for this mechanism, as related clones would continually activate each other. Instead, a mapping is required, that facilitates the activation of chains of receptors. Such chains would facilitate the ongoing maintenance (via proliferation) and processing (maturation) of information in the absence of antigen. A simple example uses the elaborated genetic scheme, where the receptor represents a complement (negative) image of an input pattern. After activation and the resultant cloning and replacement, the activated receptor set is taken as an exposure to the repertoire (given its current activity). Thus, an activated receptor becomes the focus of the receptor repertoire, which will attempt to bind to it. This binding may be attempted in a number of ways such as: (1) on the activated receptors 'receptor pattern', (2) on the conformation, or (3) on the receptors genetic composition (the first and second options provide a natural interpretation of the metaphor). If receptors are applied to a combination of the activated receptors *conformation* and *mapping* (a trinary bit string, with '#' representing 'anything') such a configuration would promote the reinforcement of inverse-blocks of contiguous antigen matching.

Integrating Feedback

[1] This is just a trivial example scheme, one may use any suitable mapping scheme off-the-shelf or of their own devising

Presently, the system is provided with information as to its general performance for each input signal (which or may not be external). Additional holistic feedback may be perceived by the system, perhaps as an analogy to the amount of tissue damage the system has currently endured (much like the danger theory of immunology). This perceived feedback may take the form of positive (reinforcement), neutral (nothing), and negative (discouragement) of the current receptors in the repertoire. The feedback may be apportioned in different ways to the receptors, for example (1) the currently active set of receptors is allocated a portion of the feedback, or (2) receptors in the repertoire are allocated feedback based on their relative recency of activation, and activation history.

***Feedback Principle**: Provides an additional information layer to receptors regarding the present holistic performance of the system*

Feedback is additional information about the systems performance, thus the system requires some way of using this information after it is allocated. Feedback provides an additional information layer that may influence the algorithm, for example:

1) *Selection of high-affinity receptors*: Those receptors that have high-affinity to a given antigen, and low feedback may not be suited to the antigen. Feedback provides a way of ignoring the specificity of receptor for antigen facilitating corrective behaviour.
2) *Cloning and maturation of selected receptors*: Feedback may provide a way of differentiating high-affinity receptors such that the amount of resources allocated (number of clones), or amount of maturation (mutation) may be tuned.
3) *Replacement of receptors*: Feedback provides an additional information layer over the repertoire at replacement time (for those receptors around during the feedback). Thus, feedback may dominant the replacement probability (when feedback is available) such that low-feedback receptors are more likely to be replaced, and the converse for high-feedback receptors.

Feedback provides system-level performance information to individual receptors, thus that receptor-antigen interactions that employ the feedback information must promote the global system interests rather than the local receptor-antigen (affinity) interests.

Spatial Structure

The repertoire may be made spatial by assigning receptors positions in a lattice structure (such as one or two-dimensional). Clones of an activated set compete and may be allocated positions in close proximity (the neighbourhood) of the activated receptors. This behaviour allows neighbourhoods of the lattice to self-organize based upon the initial distribution of receptors. The localised competition may presuppose that neighbours are genetically similar to foster the self-organization, thus replacement tournaments may be based on affinity for a current antigen, and activation history.

***Spatial Competition**: An arbitrary spatial structure (such as a toroidal lattice) for the repertoire may impose a neighbourhood-based (spatial) competition between receptors that may supplant genetic composition-based competition.*

The underlying spatial structure manifests meaning for the limitations on repertoire size, and the self-organisation qualities of receptors apportion responsibility for antigen to spatial regions. The spatial structure further reinforces concerns of take-over of receptor families as a spatial concept, and (after repertoire stabilisation) improves the efficiency of intra-cell class competition. The spatial-based competition promotes diversity as a spatial concept (on the lattice) rather than a genetic concept (through genetic comparisons), thus in addition to the efficiency offerings, may provide a more natural fit with the metaphor.

T-Cell Mediation

A second population of receptors may be introduced (T cells) that mediate which of those receptors that are activated may clone (reminiscent of the two-signal theory of mediated clonal activation). This mediation requires the insertion of an additional step in the algorithm between the selection of the activation set (Step3) and the cloning of the activation set (Step4). It may be implemented as an additional selection step, in which those receptors that are selected from the activation set may proliferate, and those that are not selected are not secondarily activated, and do not proliferate.

***Mediated Recognition**: Mediated recognition allows input signals to be remapped that applies a second level of meaning to input signals and facilitates the development of high-order information complexes (signals can be remixed and ignored)*

The two-signal approach to activation and cloning allows high-affinity input signals to be ignored. In using a second population of receptors, it allows input signals to be re-mapped and thus reinterpreted, applying meaning. The second population, in effect determines what to do with each input signal based on the context (receptors) that recognised it. High-order information complexes result from the mapping of the internal context of an input signal onto the second population, which may in turn enact an effector mechanism such as proliferation of matched first-tier (B cell) receptors.

ACKNOWLEDGMENTS

Tim Hendtlass for his patience and for providing useful feedback on drafts of this paper