Robotics and the Q-analysis of Behaviour

Pejman Iravani
p.iravani@open.ac.uk
Computing Department

Jeffrey Johnson
j.h.johnson@open.ac.uk
Design and Innovation

Lucia Rapanotti
l.rapanotti@open.ac.uk
Computing Department

*

The Open University, Milton Keynes, MK7 6AA, UK

Abstract

A new method is presented for the analysis of complex multi-agent robotic behaviours, based on observing behaviour and abstracting models on which to base control. This involves (i) building an appropriate representation of the scene using the 'best' features, (ii) classifying each scene by its features, and (iii) learning the correlation between scene classes and the actions performed in each. A novel aspect is the use of *Q-analysis* to investigate relational structure between many possible observable features, and configurations in the scene. Q-analysis helps in the selection of appropriate features and is the basis of our combinatorial classification. The methods and algorithms presented are experimentally validated using data from the RoboCup simulation league. We conclude that Q-analysis could be a powerful new approach in robot control.

1 Introduction

We refer to *behaviour analysis* as the task of constructing a model of the behaviour exhibited by an agent or multiagent system (MAS) using observations. In this paper the agents are autonomous robots working together as a team to play soccer.

Behaviour analysis starts by observing *scenes* as the interaction between agents and their environment. Scenes have sub-scenes, here called *configurations*. In the first step, the configurations are described by a set of *observed features*, for example, 'opponent to left', 'ball moving fast', 'team-mate to right', 'close to goal', or any other 'relevant' characteristic. In the second step, the configurations are classified, based on combinations of their features. Finally, the configuration classes are correlated to the observed *actions* performed by the agents. This is then used as the basis of control for the agents, effectively giving a *learned behaviour* [1, 2]. In order to analyse behaviours in this manner, the following issues need to be addressed:

1. Which features "best" describe a scene?
2. How can the features be used to classify scenes?
3. How can relationships be learned between configuration classes and the agent control actions?

The first issue, known as *feature extraction* in pattern recognition, relates to which features should and should not be part of the representation. In general, adding irrelevant features "obscures" the effects of relevant features. Thus it is desirable to find a method to include only 'relevant' features in the representation. The second issue is a general problem in classification, which relates to finding an appropriate criteria of 'similarity' between instances. The third issue is related to machine learning, in finding a method for learning from classified observed behaviours.

This paper develops a framework for behaviour analysis based on the *Methodology of Q-analysis* [3, 4]. As explained in the next section, Q-analysis is a multidimensional generalisation of network theory, able to model general *n-ary* relations between features and configurations. Through its notion of 'q-connection' it provides a graded method of classification according to shared features.

This is in marked contrast to methods of classification that map objects into multidimensional data spaces, and cluster them into components based on similarity metrics. The essential difference is that Q-analysis is very sensitive to the selection of features, and this can be exploited to detect and remove features that add little or no information.

Before starting the discussion, we establish some basic notation. A *scene* is an instantaneous observation of the system including the robots and environment. Scenes have sub-scenes associated with subsets of players that we will call *configurations*. In this paper we are concerned with selecting *features* to describe those configurations, and classifying the configurations for control purposes.

2 The Methodology of Q-analysis

2.1 Classifying Multidimensional Data

Of the large literature on classification, we can abstract two complementary approaches. To illustrate this, consider a set of objects to be classified, $A = \{a_1, a_2, ..., a_m\}$, and a set of *classificatory features*, $B = \{b_1, b_2, ..., b_n\}$. An *observation* of an object a_i consists of making two decisions for each b_j:

1. is object a_i related to feature b_j?

2. what is the strength of that relationship, weighted as a number?

Behind the first of these questions is a subtlety often lost by unquestioning use of standard classificatory techniques. Whether or not an object is related to a feature is a binary yes/no decision. For example, consider whether or not a battery-powered robot has a power source. If the battery is removed it does not (relationship). Suppose the robot has a battery fitted, but the battery is flat. Then it does have a power source (relationship), currently capable of delivering zero power (number).

This distinction makes a big difference in the application of classification techniques. When 'non-present' characteristics are included in the classification with weight zero, this impacts on the classification. In particular, it established 'similarity' between objects weakly related to a feature, and objects that are logically unrelated to that feature. In the case that the feature represents a dimension with little or no information, this creates spurious similarities.

This is one of the reasons why it is not possible to solve all problems using fully connected neural networks as classifiers. Neural networks such as the multilayer perceptron effectively map an m-dimensional input space to an n-dimensional output space. The distinction being made above corresponds to there being no connection between input x_i and neuron y_j (not related), and there being a connection for which all the input values are zero, or the weight, w_ij, being held at zero. In fact this last possibility cannot be guaranteed within the standard multilayer perceptron architecture, and the only way to achieve the required effect is to cut the link between x_i and y_j.

The practical implications of this are profound. The idea that one can throw any combination of 'data' at a network, or any other classifier, founders on combinatorial complexity. In principle the network will filter out the 'irrelevant' data by assigning low weights to their connections. In practice a network with a million inputs will never converge.

Thus, the classification of multidimensional data addressed by many clustering techniques often begs the essential question: *what are the relevant dimensions for the particular application?*

As we will show, Q-analysis is an approach that stays very close to the data. At times its sensitivity to the dimensions used can be frustrating, as can the effects of objects that dominate the structure. But often, this just reflects the nature of the system under investigation.

2.2 Similarity

In classification, *geometric* models are often used to evaluate the similarity between instances. Geometric models represent instances as points in a multidimensional coordinate space and define similarity between instances as the Euclidean distance separating them. This means that the information stored in a set of dimensions is subsumed into a distance value, and thus, it is critically important that this distance reflects accurately the relevant information in the data [5]. Figure 1 illustrates an example of geometric similarity; on the left, three robots (R1,R2,R3) are positioned with respect to a goal, on the right, two different coordinate spaces are used to represent each situation. Following the top coordinate space, robots R1 and R3 are in more similar situations than R2. On the bottom coordinate space, the angle feature (α) has been scaled differently (e.g. changed representation units), as an effect R1 and R2 are now more similar. This exemplifies the importance of finding a distance value that accurately represents similarity among instances.

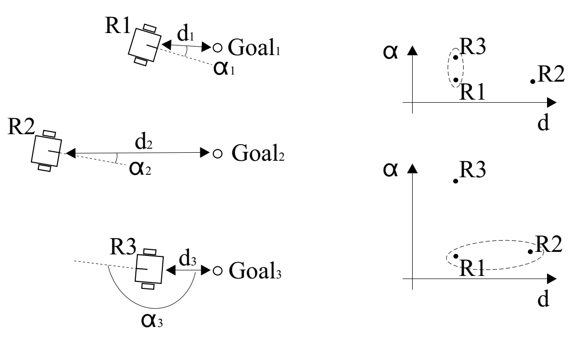

Figure 1: Mobile robots in different situations with respect to a goal and their similarity based on Euclidean distance.

2.3 Representing relations by simplices

Given the limitations of geometric models based on geometric similarity metrics, the Methodology of Q-analysis offers new insights to the concept of similarity. This analysis is especially suited for discovering *relational structure* in multidimensional data. In Q-analysis, similarity is no longer defined as a distance, but is based on structural ideas of connectivity between particular instances.

Q-analysis provides a *set-theoretic* approach to the study of relationships. In general a relation R between a set of elements, $\{x_0, x_1, ..., x_p\}$, can be considered to determine a new object called a *simplex*, denoted $\langle x_0, x_1, ..., x_p; R\ \rangle$.

Simplices can be represented by *polyhedra* in multidimensional spaces. Let the individual x_i be called *vertices*, denoted as $\langle x_i \rangle$. Then a simplex with one vertex is a point, a simplex with two vertices is a line, a simplex with three vertices is triangle, a simplex with four vertices is a tetrahedron, a simplex with five vertices is a 5-hedron, and so on. This is illustrated in Figure 2, where (a) and (b) are tetrahedra, (c) is a triangle, and (d) is line. As can be seen, a simplex with $p+1$ vertices is a p-dimensional object. Thus we refer to a simplex with $p+1$ vertices as a *p-simplex*.

We illustrate this using an example from robotics. Let $\{x_0, x_1, ..., x_p\}$ be a set of features describing a sub-scene, or configuration, in a robot soccer game. Let a configuration be denoted by c_j, so that each x_i is R-related to c_j. Then we can write the simplex associated with c_j as $\sigma(c_j) = \langle x_0, x_1, ..., x_p; R\ \rangle$.

To illustrate this consider a matrix M representing five configurations, $\{c_1, c_2, c_3, c_4, c_5\}$, related to subsets of six binary features $\{x_1, x_2, ..., x_6\}$:

$$\mathbf{M} = \begin{array}{c|cccccc} & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ \hline c_1 & 1 & 1 & 0 & 1 & 1 & 0 \\ c_2 & 0 & 1 & 1 & 1 & 0 & 1 \\ c_3 & 0 & 1 & 0 & 1 & 0 & 1 \\ c_4 & 0 & 0 & 0 & 0 & 1 & 1 \\ c_5 & 1 & 1 & 0 & 1 & 1 & 0 \end{array}$$

Table 1: a configuration-feature incidence matrix.

Each row, i, of matrix, M, can be represented by a *simplex*, $\sigma(c_i)$. Thus $\sigma(c_1)$ and $\sigma(c_5)$ are the 3-simplex or tetrahedron $\langle x_1, x_2, x_4, x_5 \rangle$ (Figure 2(a)). $\sigma(c_2)$ is also a 3-simplex, $\langle x_2, x_3, x_4, x_6 \rangle$, (Figure 2(b)). $\sigma(c_3)$ is a 2-simplex or triangle, $\langle x_2, x_4, x_6 \rangle$, (Figure 2(c)), and $\sigma(c_4)$ is a 1-simplex or line, $\langle x_5, x_6 \rangle$ (Figure 2(d)).

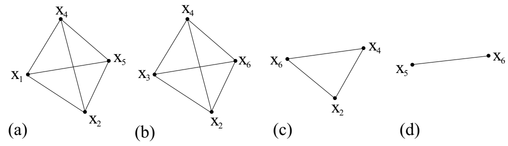

Figure 2: Some examples of different simplices.

An important idea in Q-analysis is that high dimensional simplices can be decomposed into their lower order simplices called their *faces*. For example, the simplex representing, $\sigma(c_1)$, (Figure 2a) can be decomposed into the following face simplices of dimension q:

q	#	simplices
3	1	$\sigma(c_1) = \langle x_1, x_2, x_4, x_5 \rangle$
2	4	$\langle x_1, x_2, x_4 \rangle \langle x_1, x_2, x_5 \rangle \langle x_1, x_4, x_5 \rangle \langle x_2, x_4, x_5 \rangle$
1	6	$\langle x_1, x_2 \rangle \langle x_1, x_4 \rangle \langle x_1, x_5 \rangle \langle x_2, x_4 \rangle \langle x_2, x_5 \rangle \langle x_4, x_5 \rangle$

2.4 q-nearness and structural similarity

Let the intersection of two simplices be defined to be their shared face. For example, $\sigma(c_1) \cap \sigma(c_2) = \langle x_1, x_2, x_4, x_5 \rangle \cap \langle x_2, x_3, x_4, x_6 \rangle = \langle x_2, x_4 \rangle$ (Figure 3).

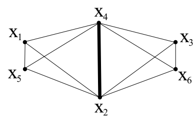

Figure 3: Q-connectivity of two simplices.

More generally, two simplices σ and σ' are said to be *q-connected* if there is a chain of pairwise p-near simplices between them, $p \geq q$.

When the relation is represented by a binary matrix, M the q-nearness can be calculated as, $MM^T - \mathbf{1}$, where, M^T, is the transpose of M and $\mathbf{1}$ is a matrix with all elements equal to 1.

	c_1	c_2	c_3	c_4	c_5
c_1	3	1	1	0	3
c_2	1	3	2	0	1
c_3	1	2	2	0	1
c_4	0	0	0	1	0
c_5	3	1	1	0	3

Table 2: A shared face matrix.

For example, the symmetric matrix above represents the q-nearness of the simplices in M given in the previous section. The diagonal of this matrix represents the dimension of each simplex, this value is also known as *q-top*.

This *shared face matrix* represents the direct connectivity of the simplices based on their shared vertices. Every p-simplex is p-near to itself. In this case, c_1 is 3-near to c_5 because they are related to the same set of vertices, and are *identical* with respect to this vertex set. On the other hand, c_1 and c_2 are both 3-simplices (tetrahedra), but they are only 1-near because they share only a 1-dimensional face.

The *conjugate matrix*, calculated using $M^T M - \mathbf{1}$, gives the number of configurations shared by pairs of features. It represents the connectivity between features, and we call it the *feature shared face* matrix.

3 Q-analysis in Behaviour Analysis

This section investigates the suitability of the Q-analysis framework for behaviour analysis, defined in the Introduction.

The underlying hypothesis is that classes of 'similar' configurations can form more abstract *concepts*. These can be associated with particular control actions that have been experienced previously by the robot as successful. Thus the classification is the means of *generalising* from particular experiences.

The study is conducted by analysing the example of robot football *passing behaviour*. The data for this come from observing 'passing configuration' from the Log-Files of the RoboCup 2003 Competition, and constructing a model of a passing behaviour. As seen earlier, to construct such model the following issues need to be addressed: (i) selecting 'relevant' feature, (ii) using these 'relevant' features to best classify 'similar' configurations, (iii) using these classes as the basis for learning behaviour.

3.1 Representing Configurations

The number of possible features that could be used to describe a configurations, or position, in a robot soccer game is enormous. Also, there is no obvious *a priori* 'best-set' of features. We constructed our set of features as follows.

Figure 4a illustrates a 'passing scene', simplified to 5 players per team rather than the 11 used in this experiment. Player, p, is the one holding the ball, while the players a_i are on the same team as p, and players b_j are the opposition team.

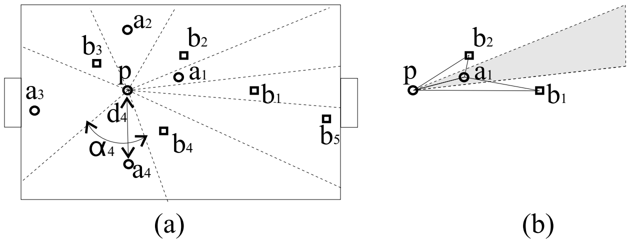

Figure 4: a) Passing scene. b) Player configuration

For a configuration, many features could be abstracted. For example, let player p have possession of the ball. Then each other player has an associated triangular area of 'controlled space' defined by p and the positions of other neighbouring players (Figure 4(a)). Then each triangle has an angle, α_i, and a distance, d_i between the player and p.

Both, α_i and d_i are continuous variables that are segmented into four intervals, 'very-small', 'small', 'big', and 'very-big', denoted as vs, s, b, vb. Two relations indicating whether or not the right and left neighbouring players are of the passer's team are also represented. We also use the feature of whether or not an opponent player is closer to the ball than a teammate. Many other features could be incorporated, but we arbitrarily select these 11 binary variables:

4 distances: d_{vs}, d_s, d_b, d_{vb}
4 angles: $\alpha_{vs}, \alpha_s, \alpha_b, \alpha_{vb}$
2 neigbours: $R_{neigh\_own\_team}, L_{neig\_own\_team}$
1 opponent closer: opp_{closer}

Thus, each configuration (Figure 4(b)) can be described by simplex with vertices subsets of these eleven features. As will be explained, these vertices themselves were the result of a selection process according to relevance.

3.2 Feature Selection using Q-analysis

Given an arbitrary set of features, which ones are the most 'relevant'? The usual approach to this question has been to *leave it to the designers intuition* to decide which are the most important features, whether new ones need to be added or if any need to be discarded. In this context, a method to evaluate features as 'relevant' or otherwise, could be helpful to human robot designers, and in the longer term, could lead into automating the feature selection process. Here we present a method of feature selection based on Q-analysis.

Let X be a configuration represented by n binary features, $X_i = \{x_1, x_2, ..., x_n\}$. For illustration it will be assumed that each configuration is classified by experience into 'passing configuration' and 'non-passing configuration'. Thus the configurations will be denoted $c_{j,passing}$ and $c_{k,non-passing}$. Some features add no discriminative information for a classification, while others give complete information.

For the simplest cases, we will define a feature to be 'distracting' when observing it adds little information in its context. For example, a feature related to *all* configurations adds no information, and a feature related to *no* configurations adds no information. The first case would apply when a thermometer always recorded $x_{temperature}$ above a safety threshold in all cases, associated with the action "sound alarm", while second case would apply when a thermometer always recorded $x_{temperature}$ as below the safety threshold associated with action "do nothing".

A feature x_i is a *perfect classifier* if all configurations of one class are are related to that feature, and no configuration from other classes is related to that feature. This last occurs when, for example, $x_{temperature}$ below a safety threshold is related to a class "do nothing", or otherwise "sound alarm".

In this paper we are concerned with discovering features that are neither distracting or perfect classifiers, but become relevant in combination with other features. In order words seek features that form the vertices of classifying simplices. These simplices can be considered to be 'concepts' at a higher level or representation [1, 2, 6].

In the next section we will illustrate a number of heuristics for selecting 'relevant' features for the case of passing and non-passing configurations. These heuristics amount to seeking feature simplices that are faces of many configuration simplices. In other words, we seek *combinations* of feature vertices that give the most powerful discrimination, and we seek features vertices that belong to many such combinations.

3.3 Q-analysis of a Games

To investigate the use of Q-analysis as a classification method, we took the Final game of RoboCup 2003 and studied the passing behaviour observed in it. Let, S, be the set of successful passing configurations (passer and receiver players belong to the same team) observed in that game.

For every pass that is made, the pitch can be divided into 21 triangular configurations, one for each of the 21 players not possessing the ball. Of these we focus on the areas of the 10 team-mates of the passing player.

Each of the ten 'team-mate' areas can be described by the eleven features defined above. Some of these are mutually exclusive, and the maximum dimension of the simplices representing the triangular configurations is $q = 4$. The pass can only be made to one of these areas, so the remaining nine become 'non-passes'.

We now apply Q-analysis on this data to study whether any structural difference emerges between the pass and non-pass simplices, both enabling us to isolate powerful classificatory simplices, and thereby relatively the powerful features that make up their vertices.

3.3.1 The Shared Face Connectivity Matrices

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_10	x_11
x_1	12	0	0	0	7	1	4	0	5	2	11
x_2	0	45	0	0	13	9	13	10	10	18	30
x_3	0	0	44	0	11	10	16	7	20	15	33
x_4	0	0	0	17	8	4	2	3	8	4	10
x_5	7	13	11	8	39	0	0	0	12	16	35
x_6	1	9	10	4	0	24	0	0	8	8	19
x_7	4	13	16	2	0	0	35	0	21	14	23
x_8	0	10	7	3	0	0	0	20	2	1	7
x_9	5	10	20	8	12	8	21	2	43	11	34
x_10	2	18	15	4	16	8	14	1	11	39	33
x_11	11	30	33	10	35	19	23	7	34	33	84

(a)

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_10	x_11
x_1	33	0	0	0	10	13	6	4	17	7	24
x_2	0	108	0	0	29	38	29	12	46	54	76
x_3	0	0	187	0	68	58	41	20	82	73	116
x_4	0	0	0	734	465	211	48	10	300	315	344
x_5	10	29	68	465	572	0	0	0	243	233	314
x_6	13	38	58	211	0	320	0	0	136	150	172
x_7	6	29	41	48	0	0	124	0	57	54	60
x_8	4	12	20	10	0	0	0	46	9	12	14
x_9	17	46	82	300	243	136	57	9	445	164	308
x_10	7	54	73	315	233	150	54	12	164	449	307
x_11	24	76	116	344	314	172	60	14	308	307	560

(b)

Figure 5: a) Pass shared face matrix, M. (b) Non-pass shared face matrix, $\tilde{M}$

Let M be the incidence matrix of the relation between the set of triangular configurations associated with passing. Let $\tilde{M}$ be the incidence matrix of the relation between the set of triangular configurations not associated with passing. Then we form the two shared face matrices $M^T M$ and $\tilde{M}^T \tilde{M}$, for the passing and non-passing configurations. Figure 5 illustrates these matrices.

The diagonal entries of the two shared face matrix give the number of passes or non-passes related to the diagonal feature. For example, out of 118 passes, x_1 is related to 12 of them while x_{11} is related to 84 of them. There are many more non-passes (1062) than passes, and x_1 is related to 33 of them while x_{11} is related to 560 of them.

Already it is clear that the vertices are responding differently to pass and non-pass configurations. For example, a comparison of the diagonals of the matrices using the 'normalised ratio' formula $(M^T M_{ii}/((\tilde{M}^T \tilde{M}_{ii})/9)$ gives the following:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
0.3	0.2	0.5	4.8	1.6	1.5	0.4	0.2	1.1	1.3	0.7

This means, for example, that x_1 is related to relatively less non-pass positions than it is pass positions (1 : 0.3), as is x_{11} (1 : 0.7). However, x_4 has a relative large ratio of (1 : 4.8)in favour of non-pass positions. This is easy to explain, since x_4 means "large distance to team mate". In fact x_4 approaches being a perfect classifier, as far as non-passes are concerned.

The further the ratios are from unity in the table above, the more discriminating are the features, x_i. As we have noted, some numbers are less that unity (better response to pass configurations) and some are above unity (better response to non-pass configurations). In combination these features may exploit these differences to give robust classification.

For example, recall that there are 118 passes and 1062 non-passes in the data. $\langle x_7 \rangle$ is related to 35 passes (30%), $\langle x_9 \rangle$ is related to 43 passes (36%), while $\langle x_7, x_9 \rangle$ is related to 21 passes (18%). In contrast, $\langle x_7 \rangle$ is related to 124 non-passes (11%), $\langle x_9 \rangle$ is related to 445 passes (42%), while $\langle x_7, x_9 \rangle$ is related to 57 passes (5%). Thus the q-nearness analysis suggests that x_7 and x_9, in combination, have more subtle discriminating power. In other words we should classify by simplices rather than vertices.

3.3.2 Star-Hub Analysis

Given any set of simplices, $\sigma_1, \sigma_2, ..., \sigma_n$, their *hub*, is the largest face of them all. Thus hub$(\sigma_1, \sigma_2, ..., \sigma_n)$= $\cap_{i=1}^{n} \sigma_i$. In the light of the comment at the last section, we should be seeking those simplices that have a relatively large hub for the 'passing' class, and a relatively small hub for the 'non-passing' class. For our set of eleven features its is possible to examine all of its 128 possible combinations ($4 \times 4 \times 2 \times 2 = 128$). Table 4 shows a summary of this star-hub analysis.

hub simplex	# pass	%	# non-passes	%
$\langle x_4, x_5, x_9, x_{10}, x_{11} \rangle$	1	1%	57	5%
$\langle x_4, x_6, x_9, x_{10}, x_{11} \rangle$	0	0%	35	3%
...	...	...	...	...
$\langle x_3, x_7, x_9, x_{10}, x_{11} \rangle$	4	3%	7	1%
$\langle x_3, x_8, x_9, x_{10}, x_{11} \rangle$	1	1%	1	0%
...	...	...	...	...
$\langle x_3, x_9, x_{10}, x_{11} \rangle$	7	6%	36	3%
$\langle x_3, x_7, x_9, x_{11} \rangle$	8	7%	13	1%
...	...	...	...	...
$\langle x_4, x_5, x_{10}, x_{11} \rangle$	2	2%	120	11%
$\langle x_5, x_9, x_{10}, x_{11} \rangle$	2	2%	74	7%
...	...	...	...	...
$\langle x_4, x_5, x_{11} \rangle$	6	5%	232	22%
$\langle x_4, x_6, x_9 \rangle$	2	2%	86	8%
$\langle x_5, x_{10}, x_{11} \rangle$	14	12%	157	15%
...	...	...	...	...
$\langle x_7, x_9, x_{11} \rangle$	15	13%	42	4%
$\langle x_3, x_7, x_9 \rangle$	12	10%	16	1%
$\langle x_2, x_5, x_{11} \rangle$	12	10%	23	2%
...	...	...	...	...
$\langle x_9, x_{11}, \rangle$	34	29%	308	29%
$\langle x_4, x_{11} \rangle$	10	8%	344	32%
$\langle x_4, x_5 \rangle$	8	7%	465	44%
...	...	...	...	...
$\langle x_2, x_{11}, \rangle$	30	25%	76	7%
$\langle x_7, x_{11} \rangle$	23	19%	60	6%
$\langle x_7, x_9 \rangle$	21	18%	57	5%
...	...	...	...	...
$\langle x_4 \rangle$	17	14%	734	69%
$\langle x_5 \rangle$	39	33%	572	54%
$\langle x_{11} \rangle$	84	71%	560	53%
...	...	...	...	...
$\langle x_2 \rangle$	45	38%	108	10%
$\langle x_3 \rangle$	44	57%	187	18%
$\langle x_8 \rangle$	20	17%	46	4%

Table 3: A selection of hubs and their frequencies.

Any discussion of Table 3 must be qualified by the observation that the frequencies are small, especially for passes. Nonetheless, distinct patterns emerge. For example, even at the relatively high dimension of $q = 4$, the $\langle x_4, x_5, x_9, x_{10}, x_{11} \rangle$ is associated with 5% of non-passes and almost no passes. At $q = 3$ there is a marked difference for $\langle x_4, x_5, x_{10}, x_{11} \rangle$ between 2% passes and 11% non-passes. At $q = 2$, for example, $\langle x_3, x_7, x_9 \rangle$ accounts for 10% of the passes and just 1% of the non-passes. At $q = 2$ $\langle x_2, x_{11} \rangle$ accounts for 25% of the passes and 7% of the non-passes. At $q = 0$ it can be seen that the vertices alone have considerable discriminating power. For example $\langle x_4 \rangle$ is a strong precursor for a non-pass (14% passes to 69% non-passes, while $\langle x_2 \rangle$ is a strong precursor for a pass (38% passes to 10% non-passes).

4 Discussion

In this paper we have identified passing configurations in robot soccer. When a player makes a pass, they have a choice of ten other members of their team to pass to. Each team member has a structure, relative to the player possessing the ball, determined as combinations of the eleven features (simplices) to characterise the game. In simple systems any one such feature or simplex would be sufficient to classify the configurations as 'pass' or 'non-pass'. In more complicated cases it is not so clear cut. We have shown that particular vertices and simplices may have strong predispositions toward one class or another.

The 'better' the vertices, the stronger will be the discriminating power. It is suggested that 'weak' vertices can be identified by belonging to few 'powerful' simplices, while 'strong' vertices can be identified by belonging to many simplices with strong discrimination between classes. Elsewhere we develop this idea towards a methodology providing heuristics for the inclusion and exclusion of features in the classification of objects determined by relational data [8].

4.1 The dynamic context

In the current context, the classification of the position is intended to underlie the decision as to which team member to pass to. In practice this is only part of the story, since the passing decision may also depend on the current (static) system state being part of a (dynamic) trajectory of states [7]. However, in this context a player may decide that the demands of the trajectory to pass the ball to particular player should be over-ridden by the strong possibility that this could lead to loss of possession given the structural position. In this circumstance the player may abort the previous plan and initiate a new trajectory.

4.2 Learning from observed populations

The examples given here have come from a single game, the 2003 RoboCup Simulation League Final, which 118 pass configurations and 1062 non-pass configurations. As the tables show the frequencies of some configurations are relatively small - probably too small for any reliable generalisation on which to base learning. Although they are outside the scope of this paper, the research prompts the following questions:

Question 1: for any given simplex used to discriminate two classes such as pass/non-pass, what is the minimum number of observations required to make the discrimination 'significant?

Question 2: can the observed frequencies for one game be added to those of another game in a meaningful way for this kind of analysis?

Question 3: can the observed frequencies of one team be added to those of another team in a meaningful way for this kind of analysis?

Question 4: what is the impact of adding features on generalisation - does combinatorial explosion reduce the frequencies to statistically insignificant levels?

These issues are studied by Iravani [8]. There he reports that the patterns observed here persist for the same team in different games, and between different teams. This suggests that data from many games can be combined to give much larger sample sizes as a more robust basis for machine learning.

Iravani also addresses the issue of machines learning over time, and even changing their tactics based on experience. For example, a particular configuration may give a good outcome initially, but then given a bad outcome when another team invokes a new strategy. In such circumstances the frequencies associated with the particular simplex(es) will 'drift' from pass to non-pass, or *vice-versa*, and the robot's behaviour will change as it learns and adapts.

References

[1] P. Iravani, 'Behaviour-based architecture for abstract control and learning', *Proc. TAROS*, 2004

[2] P. Iravani, Johnson, J.H., Rapanotti, L., 'Applications of concept grounding techniques to reduce the dimensionality in sensorory-motor space', STAIRS, 2004

[3] Atkin, R.H., *Multidimensional Man*, Penguin (Harmondsworth), 1981.

[4] Johnson, J.H., 'Some structures and notation of Q-analysis', *Environment and Planning B*, **8**, 73-86, 1981.

[5] Gordon, A. D., *Classification*, Chapman & Hall, 1999.

[6] Johnson, J., H., 'Stars, Maximal Rectangles, and Lattice: a new persepctive on Q-analysis', *International Journal of Man-Machine Studies*, **24**, 293-299, 1986.

[7] Johnson, J. H., 'Visual communication in swarms of intelligent robot agents', *Artifical Life and Robotics*, **5**, 1-9, 2001.

[8] P. Iravani, 'An architecture for multilevel learning and robotic control base on concept generation', Ph.D. Thesis, The Open University, 2004.