Quantifying Quality: The Application of Psychological Standards in Autonomous Agent Design.

Joanna Bryson†, Will Lowe‡, and Brendan McGonigle†
†Laboratory for Cognitive Neuroscience and Intelligent Systems
‡Centre for Cognitive Science
Edinburgh University
Edinburgh EH8 9JZ, UK
joannab@dai.ed.ac.uk, will@cogsci.ed.ac.uk, B.O.McGonigle@ed.ac.uk

November 17, 1997

Abstract

Although much diverse and interesting research has been carried out on mobile robots, absolute progress towards robust, reliable solutions to common problems such as navigation and goal arbitration has been slow. The problem may not have been so much a lack of ideas, but the difficulty with which to compare results and establish positive trends in research. This article suggests that part of the problem with detecting these gradients has been reliance on scientific methodologies from fields whose problems are not sufficiently related to those of mobile robotics. We propose the methodologies of *experimental psychology* as a better model.

We support our claim by:

- describing the relation between experimental psychology and robotics research,
- describing the quantitative methodology of comparative psychology and how it applies to mobile robots, and
- illustrating this methodology by using it to advance and evaluate current research on action selection in our laboratory.

We therefore endorse the use of standard psychological quantitative methods for both development and comparison of robotic behaviour. We conclude by discussing applicability of this method.

1 Introduction: Which Standard?

To prosper as a scientific research programme, mobile robotics needs a common analytic framework within which to compare theories, architectures and implementations. Although much diverse and interesting research has been carried out on mobile robots, absolute progress towards robust, reliable solutions to common problems such as navigation and goal arbitration has been slow. The problem may be so much a lack of ideas, but the difficulty with which to compare results and establish positive trends in research. This is a matter of detecting a gradient.

The problem of finding standards of assessment in mobile robotics can be understood as the problem of finding ways to quantify qualitative behavioral difference. It is not not that there are no quantitative methods available, but that there are too many non-comparable candidates. The computer scientist's interest in algorithmic complexity, processing speed or mean time to failure may have no analogue in the ethologist's test battery. Conversely, careful observation of behaviour and ecologically-grounded measures of intelligence have no reflection in the concerns of computer science. As a result, progress in the field — increasingly intelligent robots — cannot be uniquely identified. The solution must be a common framework that determines what kind of quantitative measure counts as evidence for qualitative superiority.

Submitted to the Journal of Robotics and Autonomous Systems, special issue on Quantitative Assessment of Mobile Robot Behaviour. Nehmzow, Recce and Bisset eds.

In this paper we argue that the framework of *experimental psychology* may be this common solution. Psychology also deals with the difficult problem of describing complex behaviour and comparing intelligence. Over the last century it has adopted and developed methodologies for describing results and establishing standards of comparison. We begin by describing the nature of research in experimental psychology, and contrast it to the more qualitative approach of traditional European ethology often favoured by mobile roboticists. We also explain the application and transformation of these methods and concerns to the markedly different problems of robotics. Next we detail the quantitative methods available, and provide the necessary equations and explanations for their rudimentary use.

We then give an extended example of the application of psychological methods of evaluation to both the development and the comparison of robot behaviours. We begin by detailing two mechanisms of action selection, one, Edmund, of our own development and the other, Tyrrell's Extended Rosenblatt & Payton Mechanism, available both in the literature and in on-line archives. We then proceed to describe the development of a program in our own architecture to compete with an established example of Tyrrell's, and proceed to compare and analyse the behaviour of both architectures.

Finally, we summarise both the contributions and the shortcomings of these mechanisms.

2 Experimental Psychology and Mobile Robotics

2.1 Alternative Paradigms

Since mobile robotics is an interdisciplinary field, there is a wide range of possible experimental frameworks on which to draw, each associated with a constituent sub-field. Perhaps most appealing are the simple benchmarks and direct comparisons of computer science, where an architecture may be evaluated by the number of steps it performs in a second, a disk drive by its capacity and mean time to failure, or an algorithm evaluated by its memory or time complexity with respect to the size of its input. However, even within computer science, such standards can be deceptive. The relative speed of different steps in different processors vary, thus different processors may be optimal for different tasks. Mobile robotics is explicitly about the construction of complete systems (Brooks 1991*a*, Horswill 1993). Thus the interest lies in comparing behaviours at a much higher level of abstraction.

Mobile robotics has traditionally taken inspiration from ethology. This is particularly true of research into autonomous agents and adaptive systems. See for example (Blumberg 1996) for a recent review of the role of ethology in action selection, or (Trullier, Wiener, Berthoz & Meyer in press) for a review of animal-inspired robot navigation strategies. Ethology studies animals at the level of behaviour with the aim of establishing a taxonomy of behaviour (Hinde 1970, Hinde 1982). In the European tradition this meant careful observation the animal in its natural environment, though field experiments were occasionally performed to test hypothetical behaviour decompositions (Lorenz 1973). As a result, the primitives of ethological theory are behaviours, or are mechanisms intimately connected to behaviours eg. drives and fixed action patterns (Hendriks-Jansen 1996) because these are the natural units of observation. In turn, it is this level of description at which roboticists try to emulate ethology (Correia & Steiger-Garção 1995, Blumberg 1996).

It should be noted that in the American tradition, ethology has always been an experimental science. Since the late 1960s, the distinction between ethology and comparative psychology has largely disappeared (Hayes 1994, Hinde 1982). In this paper we use the term "ethology" loosely to refer to any approach that privileges observation in a natural environment over controlled experimentation as the standard mode of examination and assessment for theories and models of behaviour.

In a sense, this emphasis is quite understandable, given that the aim of mobile robotics must be *intelligent behaviour in the environment*. By taking behaviours as the units of interest and success in real environments as the proper domain of study (Brooks 1991*a*, Steels 1994), roboticists follow traditional ethology, and can seemingly approach the problem at exactly the relevant level. However, although intelligent behaviour in an environment must be the ultimate test of a system, we suggest that the ethological aspects of mobile robotics research are part of the reason it is so difficult to compare competing theories and implementations.

2.2 The Problems of Ethology

The qualitative approach associated with the ethological approach is reflected in the processes by which mobile robot behaviour is evaluated. Much of mobile robotic and adaptive agent research is concerned with the elucidation and replication of animal behaviour (Beer 1989, Cliff 1992, Brooks & Stein 1993, Webb 1994, Maes, Matarić, Meyer, Pollack & Wilson 1996), which suggests an obvious qualitative test: We ask 'Is the behaviour lifelike? How similar is the robot's behaviour to that of the target animal?'

This approach to comparison may be adequate when there are gross behavioural differences between competing systems, or when the an important animal-like behaviour is in one robot simply absent. Similarly, the qualitative approach to behaviour development is also quite adequate when the architecture simple, and the effects of altering or removing particular design elements are predictable. But when the architecture becomes more complex, the relation between robotic behaviour and its causes becomes more indirect and we are left with a difficult inverse problem.

These problems also exist in psychology. Indeed, there is often substantial debate as to whether qualitative differences in behaviour *exist*, let alone what their underlying causes are. For example, Livesey (1986) documents the experimentation and debate on whether animal learning capacities differ only in scale or also in qualitative ability. Do monkeys learn faster than rats because they have different strategies, or only because they have more brain cells to apply to the task? These issues are addressed not by observing animals in their natural habitat, but by conducting carefully designed and controlled experiments in the laboratory. See, for example (Terrace & McGonigle 1994).

Determining which mechanism produces a behaviour simply by looking at the behaviour is an inverse problem: the inference from behaviour to mechanism is fundamentally under-constrained. That is to say that simple observation in the environment will not be sufficient to uniquely determine the causes of behaviour. This implies that we will have difficulty in comparing systems that behave similarly, which in turn implies that over the course of development, it is difficult to determine whether a design change is an advance or a hindrance.

These problems arise principally because we do not have a sufficiently operational definition of intelligence; instead we rely on our intuitions. Traditional ethology emphasises intelligence as adaptive behaviour situated in an animal's natural environment. But well-adapted specialised behaviour does not guarantee general or flexible problem-solving that is normally associated with more intelligent animals, though such animals may well be capable of displaying the same sort of behaviour (McGonigle & Chalmers forthcoming). And in particular, robots do not have a natural environment. We must therefore operationalise the idea of intelligent behaviour more carefully, in order to apply quantitative methods. This issue is described in detail in Section 3.2.

The problem of development also follows directly from the inverse problem. We can be certain that *some* change in design will improve the behaviour of a robot, but we need to know first which aspects of the current design are inadequate. Unfortunately, the inference from imperfect behaviour to a faulty element of design is also an inverse problem. Simple observation of the system in an environment does not provide enough information to uniquely determine the inadequate mechanism; the addition may be useful only in the conjunction with another change, or it may just seem to make no difference. Since we have no general way to quantify difference, the situation then reduces to that of the comparison problem.

As a result of these problems, research and testing in mobile robotics proceeds qualitatively and the development of a robot architecture depends largely on the programmer's intuition. For example, the decomposition of behaviours in a behavior-based system is entirely determined by the programmer's decomposition of the problem (Brooks 1991*a*, Wooldridge & Jennings 1995). And this decomposition is in turn determined by their intuitive grasp of the difficulties and intuitive understanding of of what solutions will be effective and tractable. Even where machine learning techniques such as genetic programming, reinforcement learning or neural networks are applied, there must be a sufficient decomposition and bias provided by the programmer to reduce the problem to something tractable. We discuss the relation between the methodology presented in this paper and the such techniques in the discussion, Section 6.

2.3 The Advantages of Experimental Psychology

Unlike field studies, a laboratory setting allows for more detailed control of inquiry and exploration of the underlying structure of behaviour. Behaviours that may appear similar in their natural context may break down differently when situated in an unnatural context. This allows the researcher to determine underlying features of the behaviour, such as triggers and assumptions. Experimental studies can support and test theories with primitives that are not necessarily visible in naturally occurring behaviours. Experimental manipulation may demonstrate that a coherent behaviour results from a mixture of disparate causal mechanisms.

There are three elements to be manipulated in the psychologists' laboratory experiments:

- **The Agent.** Working with humans who had lost their hippocampi lead to important discoveries about the dissociation between procedural and declarative knowledge (Milner, Corkin & Teuber 1968). Monitoring the electrical behaviour of brain cells has lead to evidence of explicit cells dedicated to the general control task of transitioning between elements of a sequence in monkeys (Tanji 1996).

- **The Agent's Experience.** Examples include the determination that chimpanzees who have acquired rudiments of language can form more advanced and flexible representations of causation and part-whole relations (Premack & Premack 1983).

- **The Environment.** Examples are the use of varying landmarks which resulted in the discovery that different bees in the same hive use multiple navigation strategies (Collett 1992), or presenting different surrogate mothers to an infant monkey, which lead to the discovery that being soft was more important than providing food (Harlow 1958).

At first sight, the methods and problems of experimental psychology may seem irrelevant to robotics; in contrast to the situation in psychology, we already know which internal factors combine to cause robot behaviour, since we designed them ourselves. But roboticists' complete design knowledge does not remove the utility of experimentation. As argued previously, determining the exact causes and consequences of emergent behaviour can be a difficult inverse problem. Even if we know each element of the design. we must still infer how these elements combine to produce a particular observed behaviour. Experimental methods allow us investigate qualitative difference in a constrained environment, and statistical methods allow us to quantify our confidence in any observed qualitative difference. This allows for the comparisons necessary for both evaluation and development.

2.4 The Advantages of Robots as Experimental Platforms

As experimental subjects, robots have significant advantages over their animal counterparts. This section concludes with four reasons why robots are an excellent platform for performing experiments into intelligent behaviour. The first two of these reasons relate to experimental design while the rest concern the statistical analysis of experimental results.

We can easily alter the robot.

Unlike the situation in animal research, there are few practical and no ethical problems in altering the internal structure of the robot. We may do this by adding, removing or altering physical structure on the robot, or by adding, removing or altering parts of the agent's software architecture. The former may serve to test the relevance of the agent's morphology, such as size of the bumpers or precise orientation of long-range sensors. The latter can test the utility of individual behaviours by deleting them or replacing them for comparison. Unlike lesion experiments in animals, we can be almost completely certain we have not affected any system other than the intended one, and that all changes in behaviour are a consequence of those alterations.

We can easily alter the environment.

It is perhaps not so obvious that it is easier to alter the robot's environment than to attempt the same manipulation with an animal. Since we know the nature and limitations of most sensing devices eg. radar, sonar, (Jakobi forthcoming) and we determine what sensory processing occurs, we have some idea which aspects of the environment are most salient to the robot. Clearly our knowledge of the effects of sensing is imperfect, but with rather constrained sensing devices we can make informed guesses, and test them experimentally. This is quite the opposite situation to animal research where considerable research is applied to determining what the salient landmarks and cues are.

If we know which parts of the environment are salient to a robot, then we can implement experimental controls more easily and be certain that, even though two environments look different to us, they are indistinguishable by the robot. This will speed up the process of experimentation if multiple robots are available.

We can simulate many subjects with the same robot.

Animal experiments may require many subjects to guarantee that behavioural differences are not simply chance fluctuations. With robotics, however, such work can be done with a single individual. We simulate different subjects by removing any trace of the effects of previous trials. That is, if there is a learning mechanism, if any parameters depend on experience, then we reinitialise.

With this method of creating subjects we can study some behavioural effects otherwise impossible to recreate. For example, we may be interested in whether a learning mechanism can reliably generalise knowledge acquired in domain A to another task in domain B. Rather than run many trials learning A and then transferring to B, we may take advantage of the fact that each initialisation of the robot is identical on all relevant parameters and simply save one successful instance of learning A and run multiple trials starting at the introduction to B. Clearly this type of control is impossible in animals.

We can guarantee that statistical assumptions are obeyed

Statisticians idealise the experimental situation as one where every individual in a subject population is an identical generator of data (each subject realises the same probability distribution), and the individuals are chosen randomly with replacement from an infinite conceptual population (Devore 1991, 208). If these conditions hold the data is *independent and identically distributed*. This assumption underlies the generalisation from sample to population, but is imperfectly realized in practice. However, if we use a single reinitialised robot for each 'subject' we conform exactly to the assumption: each robot is an essentially identical generator of data, since it differs only on random start up conditions, and although the probability distribution describing its behaviour may be arbitrarily complex, it is guaranteed to be the same from trial to trial. The conceptual population from which this initialisation of the robot is chosen is the class of all initialisations of the robot, which is effectively infinite.

Note, however, that a single robot may exhibit physical wear, so if multiple experimental conditions are to be attempted on this robot, then ideally experimental trials should be interleaved so that the wear affects all conditions equally.

The assumption of independent and identically distributed data is seldom perfectly satisfied in animal research. Animals may differ in ways yet to be uncovered. For example, animal language experiments using chimpanzees prior to the 1980's did not distinguish *pan troglydites* from *pan paniscus*, though the difference in linguistic ability between the two species is substantial (Savage-Rumbaugh, McDonald, Sevcik, Hopkins & Rubert 1989). If we are rigorous, we can guarantee that this will not happen when using robots.

3 Quantitative Methods for Experimental Robotics

An effective framework for comparison and analysis in robotics should explain how general qualitative judgements can be grounded in quantitative methods. It should also help deal with the inverse nature of the inference from behaviour to mechanism. In this section we consider how to use the framework of experimental psychology to assess the performance of systems involved in a similar task and to compare behaviour with and without particular design changes. The following sections will give examples of these methods from our own research and development.

This section begins with an introduction to the methods and considerations in choosing between them. Detailed descriptions of the main methods — factorial experimentation, correlation, and regression — follow.

3.1 The Use and Hazards of Quantitative Methods

The principal advantage of using an experimental methodology is that it provides a quantitative measure for determining how much confidence we should place in a result. The concept of *standard error* in analytic statistics reflects our confidence that the value of any estimated experimental parameter is within some range of the true value. In the analysis of an experiment we can choose the level of confidence we require. Since confidence is partly a function of experimental sample size, when the question is very important we may demand a high level of certainty and run many more trials. For a less crucial question, or in a limited time period we might accept more uncertainty about our estimated value.

In attempting to compare two robots experimentally we may fall into two types of error. Errors of *Type I* occur when the analysis reports a significant difference in behaviour between the two robots when they are in fact equivalent. Type I errors are errors of optimism. Errors of *Type II* occur when the analysis suggest that the two are equivalent when one is in fact superior. These are errors of pessimism. By minimizing the probability of type I errors we necessarily increase the probability of type II errors and vice versa. However, for the statistics discussed below the tradeoffs have already been tabulated, and we can decide explicitly what balance we will accept.

3.2 Operational Measures of Intelligent Behaviour

In order to treat robots experimentally, we must first decide on operational measures of intelligence. It is also necessary to fix the domain of application for these measures. For a household robot the time taken to retrieve a fixed number of soda cans might provide an adequate operational measure of intelligence, relative to the task environment. In the example from Tyrrell below (Section 4.2), a simulated agent in an artificial world is evaluated by the number of times it has intercourse, under the assumption that this is directly related to its success in passing on its genetic / evolutionary material (Tyrrell 1993). The key questions are what makes a good operational measure? and what is a good environment?

3.2.1 Choosing a Measure

Experimental considerations prompt us to find operational measures that allow a wide range of values and have a natural ordering ie. variables that can be mapped onto the natural number scale. The values for this type of variable are said to be on *interval scale*.

Interval data allows us to use the more powerful available statistical methods, eg. the t-test and analysis of variance. These statistics are much more informative than alternatives that only use information about the relative ranks of different values of the operational measure because we can assume a considerable amount about the distribution of the data in advance. In particular we can assume that the values of an interval measure of behaviour will be approximately normally distributed. Due to the Central Limit Theorem[1] normality can be guaranteed in most circumstances. The normal distribution provides a general model for judgements of difference (Section 3.4.3) and provides a straightforward account of confidence and confidence intervals (Section 3.4.2). Interval measurements are also useful for correlation (Section 3.4.4) and regression analysis (Section 3.4.5), where they also allow quantified certainty in the form of confidence intervals.

It is desirable to have an operational measure of intelligence that takes a wide range of values so that we can discriminate differences in behaviour as finely as possible. The most uninformative type of variable would be a binary measure taking 1 if a behaviour occurred and 0 if it did not. A slightly more useful, but still inadequate measure would be a frequency variable with a small range eg. a soccer score. These types of measure are little more than mnemonics for qualitative judgements of difference. We require a variable that will discriminate differences in behaviour where qualitative measures cannot.

3.2.2 Operationalising a Domain

There has been considerable interest recently in the idea of establishing benchmark tasks (See Hallam and Hayes, this volume). Benchmark tasks would correspond to the use of benchmark data sets in machine learning and statistics. Benchmarks are used to rank learning mechanisms on a small number of relevant parameters eg. generalisation error or mean run-time. Robot soccer tournaments have been suggested as one adequate test environment for mobile robots (Kitano, Asada, Kuniyoshi, Noda & Osawa 1997).

Successful footballing behaviour may well be a useful operationalisation of some aspects of intelligence eg. navigation, goal-directed activity or teamwork. However, score or tournament results alone are an unsatisfactory metric for comparison. Although the task is well-determined, whether a team wins or loses

[1] For details see any general statistics textbook eg. Devore (1991)

a small number of games is not a sufficiently rich source of information about the relative intelligence of the robots involved. This was shown particularly clearly in the recent IJCAI RoboCup competition held in Nagoya, where scores were mostly zero or one despite the fact that some teams were clearly qualitatively more intelligent than others. One team, after an equipment failure, implemented a random movement algorithm, while another team failed to move at all. The number of goals scored nor level of competition reached barely provided enough information to distinguish these cases from the competent teams that lost, therefore they do not at present constitute useful operationalisations of intelligent behaviour. The situation is, however, rather different in simulated league of robosoccer where many more games can be played and the average number of goals can be much higher.

Nevertheless, the task domain of robot soccer *could* easily yield informative quantitative measures of relative performance. To utilise robot soccer as a effective benchmark task it would be necessary to track many other variables besides the score in goals, much as is done during broadcasts of real soccer matches. Possibilities include the percentage time controlling the ball, shots on goal, and completed passes (where the ball goes from the possession of one robot to another on the same team). This should constitute a more reliable test of intelligent design and in doing so make the task more amenable to standard statistical methods.

One consequence is this change is that tracking the new metrics may be impossible to automate. Certainly from the perspective of the robot players, their perception of events may have little correlation with the external consequences of their behaviour. Deriving this information may require the same techniques as are used for deriving data in similar animal experiments: researchers observe clocked video recordings and note times and occurrences for the behaviours of interest.

3.3 Setting Up Experiments

As mentioned in Section 2.3, the use of statistical comparative methods in robotics can be applied either in comparing different systems to make claims about them, or in development. The difference between these lies primarily in how much certainty is required, and what factors will be compared. Here we briefly consider questions that need to be resolved for typical experimental set-ups for these purposes. The second half of this section explains how to calculate the relevant statistics. The following sections give demonstrations of these techniques in the development and comparison of behaviour between two architectures.

3.3.1 Straight Comparison on a Single Metric

At its most basic, comparing two systems requires only choosing the metric of comparison — something that reflects the quality by which one system is hypothesised to exceed the other. Such a comparison also requires stating a null hypothesis; what is the expected result if no difference exists? and choosing the number of trials necessary to establish the desired certainty. This latter requires running a small number of trials in advance to determine an approximate variance for the results. If during the course of the actual experimentation, the variance proves significantly larger or smaller than expected, the experiment may need to be redesigned to some extent. The use of a small, statistically insignificant experiment to indicate some idea of the extent and variance of an effect is called a *pilot study*. Psychological funding organisations normally require pilot studies before giving funding for full experimentation, due to their utility in determining both the viability and the requirements of the full experiment.

3.3.2 Comparison on Multiple Metrics

Often the comparison problem is not simply a matter of determining which of two systems are better, but also of determining *why*. Particularly in the development case, the roboticist cannot generally know in advance which parts of the design should be altered — the effects of small design changes may be lost in the general behaviour of the robot. By utilising a factorial experimental design we can assess objectively the improvement due to change on several metrics, or combinations of changes. In the simplest case we can create a experiment that pits combinations of design changes 1 and 2 against the original robot design *control* and each other in a single or multiple environments:

Design	Condition 1	Condition 2
$R_{control}$		
R_1		
R_2		
R_{1+2}		

This design allows for interactions between variables eg. we may find that although the performance of R_1 and R_2 both improve in condition 2, the improvement is more marked for R_2. Information about how design changes operate in different circumstances is vital to effective design change.

3.3.3 Comparison by Correlation

The experimental design described above reduces to multiple cases of hypothesis testing. However, design often does not amount only to adding and subtracting behaviour features, but also to determining their extent or priority in control of the robot. In this case, we want to vary a factor or factors continuously and look at their impact on overall behaviour. A simple example is the tradeoff between persistence and reactivity. In Tyrrell's world (described below), for example, the designer must address the question of how certain an animal should be of the presence of predators before allowing the desire to take evasive action to overcome persistence in courtship. An animal that is too cautious may never successfully mate, while an animal that is too persistent may be eaten before mating very many times. Similarly, a mobile robot that is too cautious in obstacle avoidance may not notice possible paths and doorways, but an insufficiently cautious robot will waste time extracting itself from collisions.

Factorial designs as in are extremely useful for establishing clear behavioural effects and determining the causal relationships between different mechanisms and behaviour. However, in the initial stages of research we may not have fully formed hypotheses about which factors are most important in determining behaviour. It is useful to have weaker, but less restrictive quantitative methods available. Correlation and regression analyses can provide information about the inter-correlations between behavioural and mechanical variables that might indicate causal relations. These methods are not as decisive as factorial experiments, but they can prove useful for suggesting why certain behaviour occurs. Regression and correlation can also be used to examine the relationship between a robot and its environment, by correlating the robots performance with continuous changes such as number or density of obstacles, availability of light and smoothness of terrain.

3.4 Quantitative Methods from Experimental Psychology

By quantitative methods we mean the machinery of *analytic statistics*, especially as used in experimental psychology. See for example (Devore 1991, Coolican 1994). We conclude this section with a short overview of some central ideas in analytic statistics — hypothesis testing, the assumption of normality and confidence intervals — and the central ideas of correlation and regression analysis.

3.4.1 Statistical Significance

In an experimental situation, the robot may be required to complete some task that we consider constitutes a good operationalization of some aspect of intelligence. Alternatively, we may simply let the robot wander as it pleases, though in this case we should also track some behavioural quantity that reflects the robot's ability. We refer to the degree to which the robot performs the task, or the predetermined behavioural quantity as B. The robots are run n times each.

For each robot we calculate the mean value of B

$$\bar{b} = \frac{\sum_{i=1}^{n} b_i}{n} \tag{1}$$

and its standard deviation

$$S = \sqrt{\frac{\sum_{i=1}^{n}(b_i - \bar{b})^2}{n}} \tag{2}$$

For these equations to hold, the values of B need to fall in an aproximately normal distribution. This is very likely the case; by the Central Limit Theorem we know that any process composed of a number of additive mechanisms will tend to produce a normal distribution in the limit of many trials (Devore 1991). Variation around the mean will be due to the many random factors in the environment over the course of trials, sensor noise, and randomized starting conditions etc.

Unfortunately, some robot behaviors may have systematic skewing. For example systematic hardware bias such as stronger motors or more sensitive long-range sensors on one side of the robot may make it more likely to turn one direction than another. Or in some cases such as the criteria in Tyrrell's Simulated Environment described below, the value of the mean may be smaller than or aproximately equal to its standard deviation, in which case there will be an sudden drop at 0 below the mean with a much longer tail in the higher values above the mean. These kinds of effects may often be compensated for statistically, for example by logging skewed data before analysis.

Even in a normally distributed sample, our value of $\bar{b}$ is only one of many possible values. Slightly different values of $\bar{b}$ would result from repeatedly running n more trials and calculating $\bar{b}$. The values of $\bar{b}$ form a distribution (which *is* guaranteed to be normal) centred around the population mean. It is this mean that is of primary interest, since it reflects the true average B value for this type of robot in this environment. We therefore treat $\bar{b}$ as an imperfect reflection of the population mean. In the limit of many trials $\bar{b}$ will converge on the population mean, but in the meantime we need to know how much faith to put in $\bar{b}$ as an estimation. We can determine how likely $\bar{b}$ is to be accurate by calculating the standard deviation of the distribution of values for $\bar{b}$ the standard error,

$$SE = \frac{S}{\sqrt{n}} \tag{3}$$

Larger values will suggest that we cannot be very certain that our estimate of $\bar{b}$ is a good estimate. This will be the case when the distribution of B values in our sample is very variable, and when the number of runs is rather low. We will use SE to modulate our confidence in our calculation of $\bar{b}$.

3.4.2 Confidence Intervals

Since the distribution of $\bar{b}$ vales is normal with standard deviation SE, it is straightforward to define a *confidence interval*. A confidence interval determines the region we are 95% (or 99%) certain that the population mean falls into. The confidence interval is centred on $\bar{b}$, since that is the best available estimate of the population mean. To calculate the interval, we make use of a canonical normal distribution[2] with mean zero and standard deviation one. 95% of the probability mass in this canonical distribution is located within 1.96 standard deviations of the mean. It is therefore 95% certain that the population mean lies on the interval $\bar{b} \pm 1.96SE$.

Clearly, as more confidence is needed (perhaps 99% is more appropriate) the length of the interval increases. This can be offset by running correspondingly more trials. The length of the interval L is a function of the SE and the required confidence

$$L = 2\ 1.96\ SE \tag{4}$$

Rearranging and substituting in the definition of SE we get

$$n = (2\ 1.96\ \frac{S}{L})^2 \tag{5}$$

which allows us to specify the length of the interval and our required level of confidence and gives the number of trials needed.

3.4.3 Hypothesis Testing and Comparison

To discover whether the robots are significantly different we will compare the difference between values of $\bar{b}$ for each robot, taking into account our confidence about their accuracy. We denote the mean value

[2] this is known as the z-score. Any normal distribution can be converted into z-scores by taking each score, subtracting the sample mean and dividing by the sample standard deviation. The resulting z-scores then reflect the canonical distribution described here.

of B for robot 1 as $\overline{b}_1$ and for robot 2 as $\overline{b}_2$. The standard errors for these estimates are denoted SE_1 and SE_2. We then calculate the standard error for the difference between b_1 and b_2

$$SE_{diff} = \sqrt{SE_1^2 + SE_2^2 - 2r_{12}SE_1SE_2} \quad (6)$$

where r_{12} is the correlation between the distribution of means $\overline{b}_1$ and $\overline{b}_2$, which we can estimate by calculating the correlation between the values of b for each robot. In our example, since the robots are of different designs and do not run at the same time, their performance should be uncorrelated, so the last term should have no effect.

The null hypothesis states that both robots will perform essentially the same way, that both $\overline{b}_1$ and $\overline{b}_2$ are actually samples from a single population. Since we are considering the difference between two means, the null hypothesis assumes that the distribution of differences centres on 0 — no difference. The distribution will be normal, so we can calculate how probable our actual difference would be given the null hypothesis. We calculate a critical ratio with b_1, b_2 and SE_{diff}

$$CriticalRatio = \frac{b_1 - b_2 - 0}{SE_{diff}} = \frac{b_1 - b_2}{SE_{diff}} \quad (7)$$

If the Critical Ratio exceeds 1.96 then it is 95% certain that the difference between b_1, b_2 is significant. In other words, the probability the two robots are the same is less than 0.05. If 99% certainty is required, 2.95 is inserted in the interval calculations above. If the Critical Ratio is still larger, then the probability is less than 0.01. In the psychological literature, 95% certainty is stated as $p < .05$, 99% as $p < .01$.

In the context of comparing robots it is also strightforward to determine the number of trials necessary to be 95% certain that the difference between $\overline{b}_1$ and $\overline{b}_2$ is within a specified confidence interval, L, such that

$$L = 2\ 1.96\ SE_{diff} \quad (8)$$

In a situation where equal numbers of trials have been run with each robot, and where there exists a rough estimate for the standard deviations of each robot's performance the number of trials necessary can be computed. Rearranging and substituting in the definition of L we get

$$n = \frac{4\ 1.96^2\ (S_1^2 + S_2^2)}{L^2} \quad (9)$$

One of the advantages of statistical methods in addressing the comparison problem is that we can treat our confidence in the result as an explicit function of the size of experiment, the number of trials and variability of behaviour. As a consequence, the equations above can be solved for each quantity. For example, the formulae for standard errors allows the experimenter to estimate the amount of confidence she should place in her results, if she can run only n trials. Conversely, she may calculate how many trials will be necessary to demonstrate difference with some predefined confidence. To discover how many trials will be necessary we compute

We can apply substantially the same reasoning to the problem of testing a single robot in different circumstances, or with two possible design changes. In the latter case we use the design described above. We may run any number of slightly redesigned robots with the same task.

We may also examine the effects of the environment on a particular design by running a single robot in different circumstances. If, for example, we believe that certain sensing routines are more effective in particular environments, we may test this conclusively. The principal difference in this case is that the correlation term in the formula for SE_{diff} (Equation 6) will be important, since we are measuring the robot's performance in two circumstances.

We may also wish to compare more than 2 environments or robots in the same experiment. For this type of experimental design we may use analysis of variance to calculate significant differences in behaviour for arbitrary combinations of robots and environments. The statistical computations are much the same, though rather more involved. However any general statistical software will be capable of analysing this type of design. The analysis in Section 5.3 was done with Perlman's Pipestat package (Perlman & Horan 1986), which is contains most common statistical measures, is freely available and runs from the command line.

3.4.4 Correlation

The measures described above all require the careful specification of a hypothesis and a relevant behavioural quantity that will reflect intelligence. Under these specifications the experiments can be analyses with considerable certainty. However, in many situations we will not have such developed ideas about which behavioural quantities are relevant, or which environmental factors are important. We would like to use statistical methods in a more exploratory way to suggest variables that might be predictive or explanatory of the robot's behaviour. The methods described below all rest on the idea of correlations. In simple correlation analysis we may examine to what extent two variables, for example the number of robots vs. the time a collective task takes to complete (Fontan & Mataric 1996), tend to covary. In regression analysis we ask to what extent a single variable may be predicted by knowing the distribution of others.

Two variables are said to be correlated if they are linearly related, if they increase or decrease together. Correlations are most useful for suggesting causal relations and possible explanations. Although correlation does *not* necessarily imply any causal relationship between the two variables, if there *is* a reliable relationship it should show up as a stronger correlation. On the other hand, the correlation may indicate that both measures are reflecting some unidentified causal factor. In this robotics is at an advantage over psychology, in that there are generally a limited and known number of factors that might cause any particular behaviour. Thus finding a correlation between that behaviour's expression and the activation of a particular software module or modules can be a strong indicator of how that behaviour is being generated.

We can measure correlation for n instances of x and y by computing r

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{10}$$

r is invariant to the units of x and y. It varies between 1 for a perfect positive correlation to -1 for a perfect negative correlation.

Although we cannot consider them here, there are methods by which to quantify our certainty in any value of r which rest on the same principles as those described above.

3.4.5 Regression

In the simplest case of regression, we seek a linear relation between variables x and y such that x is a predictor for y. We assume that the correlation between the two variable is not perfect and postulate a noise model for the prediction of y. Specifically, we try to fit the regression equation

$$Y = \beta_0 + \beta_1 x + \epsilon \tag{11}$$

This is an equation for a straight line ($y = mx + c$) with a noise model ϵ that is normally distributed with a mean of zero and a standard deviation of σ_ϵ.

The parameters are fitted by minimising the sum of squared errors over n data points

$$SSE = \sum (y - (\beta_0 + \beta_1 x))^2 \tag{12}$$

and estimating the standard deviation of the noise model by

$$\sigma_\epsilon = \frac{SSE}{n-2} \tag{13}$$

We may add any number of additional variables to the equation if we consider that may help predict the value of y. Each additional variable is supplied with a corresponding β value and the parameters are refitted. This process may be familiar from the field of neural networks; it is equivalent to training a one layer perceptron with one weight (β_1) and one bias (β_0) with added noise.

The goal of regression analysis is to evaluate the extent to which variables $x_{1..i}$ are effective predictors of the value of y. We quantify this by computing the total sum of squares for the data

$$SST = \sum (y - \overline{y}) \tag{14}$$

The total sum of squares is a measure of the dispersion of the data points from a horizontal line (or hyperplane if there is more than one x) defined by $y = \overline{y}$. This line corresponds a model that postulates no correlation between x and y. From this 'null hypothesis' we calculate the coefficient of determination

$$r^2 = 1 - \frac{SSE}{SST} \tag{15}$$

The quantity r^2 denotes the proportion of variation in the values of y that can be explained by assuming that they have a noisy linear relation values of x. It is in fact the same r that is described by the correlation equations in the previous section.

Regression analysis is widely used in experimental psychology because it suggests possible causes x for a variable of interest y (Coolican 1994). The model is useful because it is easy to interpret due to its linearity. Non-linear variants exist, although the interpretation of such models becomes increasingly difficult as they become more similar to distributed neural networks.

4 Examples from Active Research

In the previous sections we have recommended the application of experimental psychological method to the field of mobile robotics, then given a brief tutorial on these methods. This section illustrates the previous sections by documenting the testing of a mobile robot architecture against an action selection mechanism developed in simulation. The robot architecture is Edmund (Singh in press), a behavior-based architecture built to address the problem of programming complex behaviour in a reactive architecture. The established action selection mechanism is the Extended Rosenblatt & Payton Mechanism (ERP) developed by Tyrrell in his thesis (Tyrrell 1993). Tyrrell compares five action selection mechanisms, including drive theory, Lorenz's Psycho-hydraulic theory and Maes's Spreading Activation Mechanism, and finds an extension of Rosenblatt & Payton (Rosenblatt & Payton 1989) to be the best option. He also argues that this solution is nearly optimal.

This section begins by briefly describing Edmund and our attempts to test and validate the architecture to date. Next we describe Tyrrell's ERP and the environment in which it was developed. We then document our use of regressions and correlations in developing a program under Edmund in Tyrrell's environment, and finally use factorial experimentation to compare the results of the two architectures. The final section will discuss the strengths and the limitations of the method with reference to these results.

4.1 The Edmund Architecture

Edmund is intended to extend behavior-based approaches to artificial intelligence, such as subsumption architecture (Brooks 1991*b*) to allow for behaviour of greater complexity. Subsumption architecture stipulates that intelligence should be decomposed into individual behaviours, each of which controls its own sensing and activation. All behaviours run continuously in parallel. Whether a behaviour is expressed in the control of the robot is determined partly by the behaviour itself, which can recognise situations in which it should be active, and partly by intricate interactions with other behaviours. Behaviours are allowed to inhibit and change each other's input, and also to suppress each others output. Behaviours are organized in carefully crafted networks that specify these interactions. Groups of behaviours for pursuing the same goal are implemented as a unit called a level. Levels of additional, higher goals are added in a linear hierarchy. The behaviours in higher levels may inhibit or suppress lower levels' behaviours, but lower levels are intended to remain "unaware" of higher. The lower levels' behaviours should be unchanged from their internal perspective.

Subsumption architecture works well in situations where there is a small number of non-conflicting goals that require behaviours that are fairly homogeneous across the entire of the robot's intended niche (Horswill 1993, Bryson 1995). However, it is difficult to specify the performance of different behaviours in the same physical environment, such as when a sequence of behaviours needs to be conducted, or when the robot is selectively pursuing one of multiple possible goals (Connell 1990). We address this problem in Edmund by creating a hierarchical control system for triggering the expression of behaviours. This control is based on two units: the *competence* and the *action pattern*. An action pattern is a sequence of behaviours that will execute in turn unless interrupted by a radical failure or a higher priority goal. A competence is a grouping of action patterns and other competences which express a behaviour consistent with a particular goal. The elements of a competence are prioritised so that if multiple elements are able to run, the element closest to achieving the goal is selected. Competences are similar to reactive plans (Agre & Chapman 1988), or triangle tables (Nilsson 1984, Nilsson 1994). The idea of using hierarchies and sequences to improve behavior-based control is also widespread (Firby 1987, McGonigle 1991, Gat 1991, Tyrrell 1993, Correia & Steiger-Garção 1995, Levison 1996), though still controversial (Hendriks-Jansen 1996, Maes 1991).

Edmund differs from other behavior-based architectures in two ways. First, it completely separates the issue of flow-of-control from flow-of-information. The only information available in the control hierarchies is recent control and decision state. All perception and sensory-based memory is present in the behaviours themselves. The flow of information between behaviours is not restricted by the control architecture, but rather specified within the *behaviour library*, customised for the particular robot platform. Adaptation, both in terms of short term perceptual processes and long term "lifetime learning," occur in the behaviour library. The variable state needed for learning is stored in the behaviours that use it.

The second difference distinguishes Edmund from other behavior-based architectures that exploit plan-like structures for organizing complexity, such as RAPs (Firby 1987) and the other hierarchy-based architectures mentioned above. Although hierarchical, the control architecture supports both reactivity and arbitrary chains and loops of behaviour. This is achieved through several methods:

- Each element of a competence is prompted to perform by its own perceptual requirements. There is no error passing or assumption of success for an operation. An action sequence will complete regardless of the success of its elements, the result is checked at the competence level. Infinite loops are prevented and alternative strategies enabled by a system of habituation.

- The top level of the hierarchy consists of special competences called *drives*. Drives execute all their elements in course-grained pseudo-parallel with best-attempt scheduling for the frequency of attention to the varying elements. This allows for the constant monitoring of important sense modalities or internal states and the possibility of interruption for urgent or opportunistic conditions.

- No stack is maintained as control traverses the hierarchy of competences. When a competence prompts the activation of one of its constituent competences, the new competence takes its place in receiving the attention of the parent drive. This enables infinite loops in behaviour, such as oscillating between active exploration and grooming/memorising phases of behaviour in rats exploring new enclosures (Wilson & McNaughton 1994). Drives maintain a record of there initial elements, so that if a competence chain completes, the chain may be started again from the beginning. If nothing has changed in the world, control will return to the parent of the competence that just completed which allows it to complete, while if something relevant has changed control may switch to a more appropriate competence or action pattern.

4.2 Tyrrell's Environment and the Extended Rosenblatt & Payton Mechanism

The first step towards assessing the success of an architecture intended to support the development of complex behaviour was a definition of complex behaviour. Unfortunately, no universally accepted definition of the term exists. Tyrrell, in his dissertation (Tyrrell 1993), creates a task in a simulated environment which requires an architecture to manage a large number of conflicting goals. He defines an environment in which a small omnivorous animal is required to survive and breed. He defines six "subproblems" for the animal to solve.

1. Finding food and water. There are three types of food and three forms of nutrition satisfied in varying degrees by the different food types.

2. Escaping predators. There are both feline and avian predators, which move at different speeds and have differing powers and problems of perception.

3. Avoiding hazards. Benign dangers in the environment include ungulates, cliffs, poisonous food and water, and temperatures varying above and below ideal during the period of the day. The environment also provides various forms of shelter including trees, grass, and a den.

4. Grooming. Grooming is necessary for homeostatic temperature control and general health.

5. Mating. The animal is assumed to be male, thus its problem is to find, court and inseminate mates. While it is hazardous to try to inseminate unreceptive mates, rearing of offspring is considered to be the problem of the mate.

6. Sleeping at home. The animal is essentially blind at night; sleeping in its den is important for avoiding predators and for avoiding being trodden on by ungulates. It is also necessary for maintaining temperature while conserving energy.

These problems vary along several metrics: homeostatic vs. non-homeostatic, dependency on external vs. internal stimuli, periodicity, continual vs. occasional expression, degree of urgency and finally, whether it is prescriptive or proscriptive with regard to particular actions. In addition, the environment is dynamic and sensing and action uncertain. Perception in particular is extremely limited and severely corrupted with noise; the animal usually misperceives anything not immediately next to it, unless it has chosen to expose itself by looking around in an uncovered area. Finally, action is also uncertain, particularly if the animal is unwell.

Tyrrell separates the problems of learning and navigation from the problem of action selection by providing these elements in his simulation. Thus the animal has available as primitives a direction in which it thinks it remembers its home or recently observed food and water. The animal's sense of location with respect to its den decays over time and distance, thus keeping track of its bearing is a part of the "sleeping at home" sub-problem.

Tyrrell checks the independence of his results from these parameters by running all his models in four worlds. Besides the standard model he first designed, there are three variant worlds. According to the thesis (p. 162) these changes are:

1. Perception is altered by affecting the animal's visibility according to the time of day, the amount of vegetation, and the animal's own activity.

2. The noise variance for navigational information is tripled, and the size of the remembered map is halved.

3. Motor control is compromised by making it more probable that incorrect actions are be taken, and by changing the conspicuousness and energy consumption of various actions.

Tyrrell implemented and tested four theories of action selection, often extending and providing necessary and obvious modifications to the theories in the attempt to make them operate on a real system. The first was drive theory ("simple motivational competition") as described by Hull, the second combined Lorenz's two "Psycho-Hydraulic" models, the third implemented Maes' spreading activation networks, and the fourth extended Rosenblatt and Payton's "connectionist, hierarchical, feed-forward network." Tyrrell recommends a modified version of the last theory, so Edmund was tested against this model[3]

The Rosenblatt & Payton model had, like Edmund, been developed and proven as a robot architecture (Rosenblatt & Payton 1989). The main feature of their architecture is what Tyrrell terms a "free-flow hierarchy". Nodes in the hierarchy receive activation from internal and external stimuli and higher elements of the hierarchy. They pass energy to their children. The ultimate leaf nodes of the tree are actions, which are activated on a winner-take-all basis. What differentiates the model from the drive model or other normal hierarchies is that no decisions are made until the leaf or action nodes[4]. This allows for *compromise candidate* behaviours to be selected, that is, behaviours that satisfy multiple drives. The model is distinguished from Maes's network by its lack of preconditions (in the ethology literature, releasers), and the fact the energy flows in only one direction.

An example of a section of Tyrrell's ERP hierarchy for his simulation is shown in Figure 4.2. The internally gauged distance from the den and the externally cued night proximity both feed the high-level node "Sleep in Den", which in turn sends positive energy to nodes like "Approach Remembered Den" and "Approach Perceived Den" and "Sleep." For the remembered den temporal and uncertainty penalties send negative activation to the node as well. Many nodes send weights to the various possible directions

[3]Interestingly, Lorenz's explicitly ethological Drive theory was placed a fairly close second, whereas the only previously fully implemented, simulation-based AI architecture, Maes's spreading activation network, was by far the poorest performer.

[4]This approach has been taken more recently by Blumberg (1996).

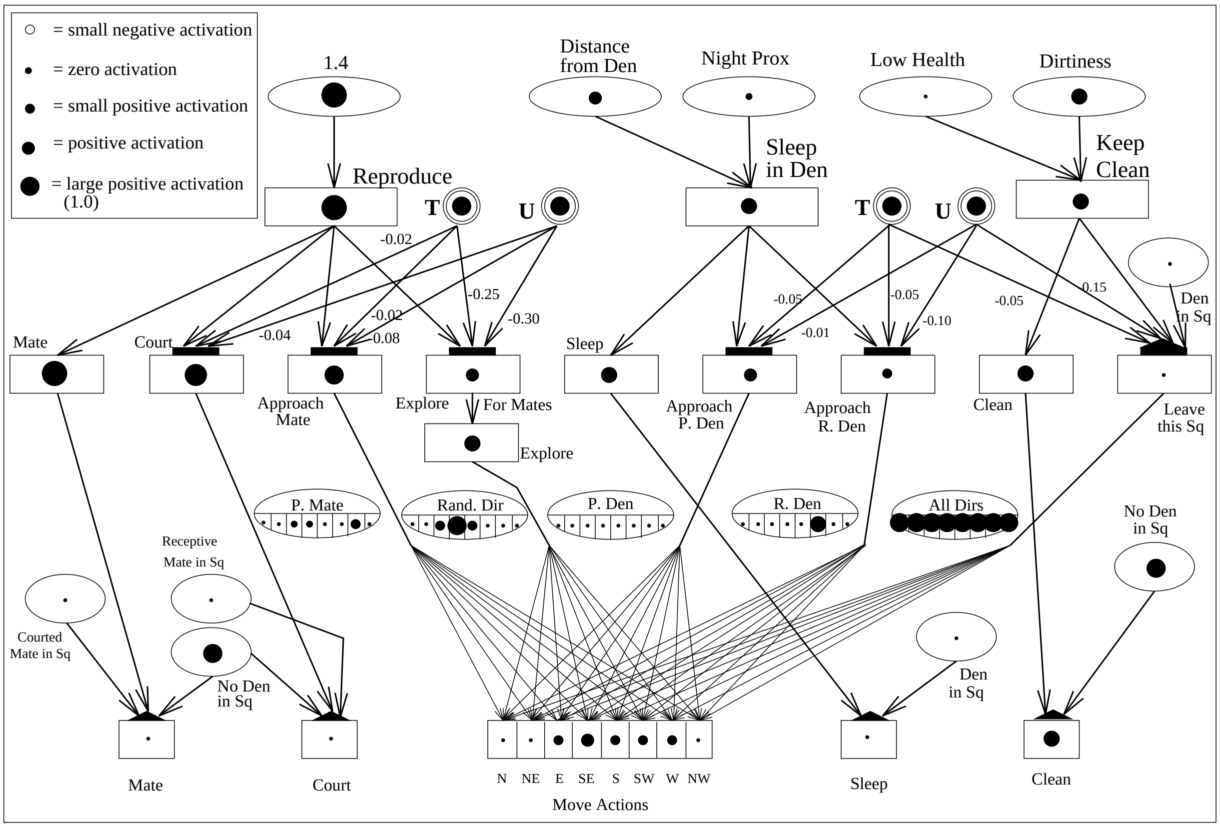

Figure 1: A fraction of the Extended Rosenblatt & Payton action selection network Tyrrell developed for controlling an animal in his simulated environment. This shows the interaction of the hierarchies associated with reproduction, staying clean, and sleeping in the den. **T** and **U** are temporal and uncertainty penalties for actions that take time or pursue a goal uncertain to succeed. "P." stands for perceived, and "r." for remembered. From Tyrrell (1993), page 247.

the animal can move in; care must be taken that motion doesn't always triumph over consummatory acts with few sources of input.

Tyrrell's main extensions to Rosenblatt & Payton's original architecture are to add penalties for temporal delays and uncertainty of rewards for items that are distant or based on poor memory or perception. The architecture then tends to exploit certain local reward rather than choosing to pursue uncertain but potentially larger goals. All perception in the simulation is real valued, so food sources of the same type in different locations will have different effects on the levels of activation energy to the network. He also added explicit penalties to appetitive vs. consummatory actions. Finally, Tyrrell created a new rule for combining the inputs to a single node that explicitly distinguished the maximum of inputs adding positive energy to a particular node and the minimum of inputs opposing the node[5].

4.3 Previous Evaluations of the Architectures

Previous to the work documented below, the evaluations of Edmund had been entirely qualitative. As a result of the claim that Edmund allowed for complex behaviour, its first test was in a domain often considered too complex for reactive systems: blocks world assembly with perception provided by a visual routine processor (Horswill 1995). The complexity of this domain lies in the number of correctly sequenced steps required to reach a solution. At the same time, it requires reactiveness, since neither the starting

[5]Humphrys (1996) discusses various action-selection mechanisms based on considering positive and negative opinions of multiple goal-embodying behaviours.

configuration nor the results of actions are predictable. It was relatively easy to construct program scripts in Edmund that replicated other work in this domain. Whitehead (1992) uses reinforcement learning to build a global behaviour for finding and grasping a green block buried in a collection of blocks. The reinforcement learning technique could not cope, however, with the adequate state variables necessary to complete the intended task of placing a red block on a green block. Edmund's hierarchical control structure allows it to focus control attention on a particular aspect of a task and thus not require additional perceptual variable state. We also replicated the "copy demo", where the goal is to duplicate an observed stack of blocks. These tasks are well beyond simple reactive systems, which have no means of selecting coherent, persistent behaviour while the environment stays static. However, they would be fairly trivial to solve using a modern planner with reactive elements to deal with the uncertainty of action and perception (Georgeff & Lansky 1987, Chapman 1989).

The complexity most often mentioned in mobile robotics literature is the complexity of the "real world;" the complexity of noisy sensors, irregular shapes, uncertain location and dynamic environments. Some critics of hierarchical structures claim that they result in sensory bottlenecks and missed opportunities making them impractical for embodied systems pursuing multiple goals (Hendriks-Jansen 1996, Maes 1991). Due to time pressures, Edmund's blocks world implementation was never connected to a real robotic arm, though it was connected to a camera-based visual-routine processor. However, the next testbed for Edmund was a real robot — a commercial Nomad 200. The robot is radially symmetric with 16 sonar and infrared sensors and 20 bump sensors arrayed around its trunk. Information from these noisy and unreliable sensors needs to be fused with each other and across time. In addition, there is information from bump sensors, clocks and motor monitors to give information about the location of obstacles and trajectories. We implemented behaviours that allow the robot to wander or be shepherded through an office, including through doorways little larger than itself. It currently maintains the separate goals of avoiding collisions, attending to bumps when they occur, attending to battery level, and performing rudimentary mapping and navigation.

Although these tasks to some extent validate Edmund by demonstrating it in a second complex domain, there are many mobile robot architectures for performing these kinds of tasks. In an attempt to demonstrate behaviour of higher complexity, we are working on producing phased behaviours, so that the robot can display a larger number of qualitatively different behaviours in smooth transition. This work involves balancing goals, being able to both lose and recover "stacked" goals when appropriate. Such behaviour would be relatively novel in a mobile robot. However, novelty is neither always possible nor desirable in comparing the relative merits of robot architectures.

Tyrrell's Extended Rosenblatt & Payton architecture does not appear to have been tested in domains other than the one in which it was designed. However, in that domain it was tested thoroughly against several of the dominant theories of action selection using statistical methods. Because the simulated environment offers the opportunity to run a sufficient number of trials for statistically reliable comparisons of even relatively small differences, and because constructing even elaborate control structures is simpler, at least under Edmund, than creating valid and useful perceptual routines for an architecture on a robotic platform, we chose to compare Edmund and Tyrrell's ERP in Tyrrell's Simulated Environment. The following section details our development approach and results, illustrating the methodology of the previous sections.

5 Comparing Edmund to the Extended Rosenblatt & Payton Architecture

5.1 Understanding the Problem

Our first step was to compile Tyrrell's simulation and the ERP, and to test that our results for his architecture matched his. We decided to test only to a 95% certainty of being within 10% of the correct result, that is to run only 245 simulations. We actually did this twice; our results were:

Run	Mean	Std. Dev.	SE
1	8.94	8.33	0.54
2	8.14	7.78	0.25
Thesis	8.31	6.63	0.16

Although these values do overlap each other's ranges within the required ten percent interval (and Tyrrell's reported result is also situated within that overlap) this shows the high degree of variability still possible, despite the small standard errors:

Working from the assumption that Tyrrell was correct in his analysis and had created a near-optimal architecture for exploiting his environment, we determined that the best opportunity to exceed his performance would be to exploit Edmund's superiority at expressing ordering and persistence. Unfortunately, very little in the simulated environment requires persistence. Mating is the only action which requires multiple steps: getting in the same square with a receptive female, courting and then consummating. Because the potentially elaborate courtship ritual had been reduced to a single act, and the receptive state of the female may be sensed, each of these elements may actually be treated as a separate reactive task.

We decided the simplest way to test persistence would be to substantially reduce the food supply. In the original test environment, food and water are both plentiful, and likely to be available nearby to any given square. Decreasing the probability of local availability of food would presumably increase the probability of dithering, which should lead to decreased efficiency and therefore genetic fitness.

In fact, however, even severe reduction of food had relatively little impact on the performance of Tyrrell's animals. Eating turns out not to be a particularly important part of life, because average life expectancy is fairly short, food is plentiful, and energy consumption is low. Consequently, changing the range of possible initial amounts of the various food types as follows:

Food	Original	Sparse
fruit	[50, 81]	[5, 35]
cereal	[45, 26]	[2, 27]
prey	[25, 36]	[2, 15]

where the range is square-bracketed, resulted in only a slight loss of fitness:

Run	Mean	Std. Dev.	SE
Sparse	6.09	6.95	0.09

Linear regression across the availability of food, water and mates showed only a few barely perceptible dependencies:

Variable	gen. fit.	water	cereal	cover	fruit	mates	prey	lifespan
gen. fit.	1.000							
water	0.026	1.000						
cereal	0.083*	-0.018	1.000					
cover	0.046	0.013	0.025	1.000				
fruit	0.206**	-0.039	-0.008	0.023	1.000			
mates	0.038	-0.009	-0.012	-0.032	0.057	1.000		
prey	0.034	0.003	0.034	0.019	0.061	-0.016	1.0000	
lifespan	0.912**	0.022	0.081	0.070	0.231	-0.010	0.018	1.000

We follow the standard notation in psychological literature, where one asterisk denotes a correlation of significance at the 0.05 level ($p < .05$), and two asterisks indicates $p < .01$. The number of initial mates is not significant because mates wander in and out of the animal's territory continuously, so their initial numbers are not particularly relevant.

5.2 Development

The development of an Edmund architecture for the Tyrrell Simulated Environment happened in two phases. The first was a period of rapid development without testing of initial versions of the primitive senses and actions for the architecture. This phase was the culmination of a period spent reading the thesis and testing the Tyrrell architecture (about a week) but took approximately 9 hours of coding time, most of which was devoted to finding the names and standard values of the necessary sensing parameters in Tyrrell's code.

The initial actions were: sleep, hold_still, pick_safe_dir (for fleeing), go_fast, go, look_around, pick_home_dir, copulate, strut, pick_mate_dir (for pursuing a mate), exploit_resource (take advantage of anything in this cell that the animal needs) and pick_dir (the general case for direction selection.) Senses were:

at_home, covered, see_predator, late (in the day), day_time, fit_to_mate, courted_mate_here, mate_here, perceive_mate, and needed_resource_available. It was clear from this selection that the main problems to be addressed individually would be fleeing predators, pursuing mates, and getting home before dusk. Other needs could be met opportunistically when nothing else was troubling the animal.

The next phase of development involved testing which behaviours were actually useful and in what order of priority, and tuning the behaviour of the primitives. Tyrrell (1993) states that roughly the same amount of time was spent tweaking parameters for each of the architectures he tested in his world, but he does not specify how long that was. In any case, these times will have varying significance for individual programmers depending on their ability and experience with the task domain, and between programmers depending on relative abilities and familiarity with the language. However, for a ball park figure, approximately 18 additional hours were spent in tuning performance. This period was spread over 10 days, since many of the analyses took 12-18 hours to run.

We began with the most basic program that could generate a genetic fitness measure – one that only pursued mating. Thus the drive "life" had a single competence "mate", which consisted of:

life (D)	mate (C)	inseminate	(courted_mate_here t) copulate
		court	(mate_here t) strut
		pursue	pick_dir_mate go

pick_dir_mate might choose to hold still if a mate was present but in a dangerous square (where there was a cliff or a predator, for example.)

This simple program performed fairly well, though on average it lived under half a day, it had a genetic fitness of slightly more than 2, immediately outperforming Tyrrell's best implementation of Maes's Spreading activation network. Other obvious behaviours were added gradually, testing the new additions and ordering on approximately 800 runs to determine if there were significant changes. If there were, then simplicity was favoured. After this phase of development the components of life in order of priority were:

life (D)	flee (C) (sniff_predator t)	run_away	(see_predator t) pick_safe_dir go_fast
		freeze	(see_predator t) (covered t) hold_still
		look_for_predator	observe_predator
	mate (C) (sniff_mate t)	*as above*	
	home *1 in 5*	(late t) (at_home nil) pick_dir_home go	
	check *1 in 5*	look_around	
	exploit (C) (day_time t)	use_resource	(needed_resource_avail t) exploit_resource
		leave	pick_dir go
	sleep_at_home	(at_home t) (day_time nil) sleep	

Where homing in the evening and looking around for predators took priority intermittently, but habituated to allow other, lower priority behaviours to express themselves immediately after they had fired.

At this point, the genetic fitness of the Edmund animal was approximately 5. The next phase of development was to examine the animal's performance step-by-step in the graphic user interface. This rapidly[6] gave information that was not available simply from the genetic fitness and the cause of death; the only information otherwise available:

- The animal seemed too concerned with staying in shelter and failed to explore much unless chasing mates. This was altered by reducing the desire for shelter in pick_dir.
- The animal was often killed by predator1's while freezing. Consequently we added a prohibition to the freezing action pattern that prevented freezing in the presence of predator1.
- Once the animal was less concerned with staying in shelter, it sometimes explored to such a degree that it lost track of its bearing entirely and could not recover by nightfall even though it began looking for its den with plenty of time. Consequently we added the action pattern:

triangulate	(lost t) pick_dir_home go

  which had the same priority as homing above, but was active all day and did not habituate. Thus, once the animal becomes slightly lost, it goes towards its den until it recognises where it is (unless it happens across a mate or a predator.)
- The animal often had unforced accidents with irrelevant animals and cliffs, particularly while it was pursuing mates. We altered pick_mate_dir to increase the avoidance of these problems.

Our observations lead to some ambiguities. Clearly there were tradeoffs for setting the animal's parameters. Because the animal had very erratic perception, it frequently hallucinated a predator nearby. Attending to any trace of a predator disrupted behaviour coherence sufficiently to interfere with mating, while paying too little attention to predators reduced mating opportunities by reducing the expected lifespan. The fleeing competence allowed for two thresholds, sniff_predator, above which the animal should take a moment to observe the predator and resolve the ambiguity, and see_predator, above which it should take evasive action. But what were the optimal values for these parameters, and were they the same for both types of predators? We varied four values (the two values for each predator) and perform regression across the results. We also decided to vary the extent to which shelter was sought and dangerous places avoided.

Unfortunately, linear regression was uninformative for the values of the predator variables, because, as described above, the expected relationship was at least quadratic, not linear. Thus we executed a larger number of trials for our data (4400) and tried two different sorts of analysis on that data. On the assumption that the variables were independent, we derived optimal values for each of the parameters via non-linear (quadratic) regression. Unfortunately, the results had high variance and may have been multi-modal, thus indicating that the variables were in fact highly interdependent. Consequently, our second strategy was simply to select the parameters of the most successful runs. Our final winning animal was in fact a product of this latter approach. Its overall genetic fitness was approximately 6.6, whereas the animals that reflected peaks on the independent curves performed as poorly as 5.0 in the worst case.

5.3 Comparison

Most of the development of the Edmund animal was done in the version of the world with sparser food described in Section 5.1. We never exceeded the overall mean value achieved with the Extended Rosenblatt & Payton model. However, in the course of analysing the results we discovered that this was the consequence of one particular test situation. The Extended Rosenblatt & Payton model seemed to perform particularly well in the variant world 3. As discussed in Section 4.2, this world is supposed to have different, implicitly more difficult, motor models. However, examination of the code proved that incorrect action selection was only increased when health was below ten percent, in which case the animal is quite likely to die soon anyway. More importantly, energy consumption was drastically reduced — mating required only half as much energy as usual and looking around, moving and running required only a third as much energy. Conspicuousness had also been changed significantly, with some motions, such as pouncing, being more conspicuous and others such as mating, being less so. The ERP was able

[6] In approximately 50 animal life-times, which took about 5 hours to observe while making corrections.

to significantly exploit these changes, with the life-expectancy in days of the animal raising by a factor of nearly three for this world:

Worlds	life expectancy (s.d.)
Standard	1.31 (2.03)
Var. 1	0.91 (0.94)
Var. 2	1.26 (1.32)
non 3 (comb.)	1.16 (1.25)
Var. 3	3.75 (2.60)

In fact, the critical ratio in comparing variant world 3 with its nearest equivalent, the standard world, is 2.94, or in other words they are significantly different with a $p < .01$. The CR between variant 3 and the other three worlds combined is 7.52. As shown in the regressions (both above and below), life expectancy is very highly correlated with genetic fitness.

In the sparse world, Edmund actually exceeded the ERP if world 3 was not taken into account:

World	Edmund	Tyrrell	Critical Ratio
Standard	6.12 (6.00)	4.77 (4.75)	7.06**
Var. 1	2.77 (2.93)	2.46 (2.56)	3.12**
Var. 2	6.20 (6.25)	4.56 (4.76)	8.40**
Var. 3	7.66 (7.50)	12.53 (9.06)	17.09**

For each architecture, the mean genetic fitness is followed by its standard deviation in parentheses.

Note that Edmund succeeds even when perception was reduced, in variant world 1. This is surprising, since perceptual ambiguity should be more damaging to Edmund's more ordered action selection. The ERP is essentially always seeking the optimal action, thus perceptual ambiguity affects all behaviour equally, whereas Edmund has different active goals and concerns. Edmund's advantage as an architecture lies in reducing the complexity of design for the programmer and focusing attention for the animal: actions that require an ordering can be set to one explicitly. The deficit in the third world shows that Edmund's animal probably wastes too much time and/or exposure to predators in stalking food. This is borne out by the fact that Edmund is soundly beaten when food is back to normal availability.

World	Edmund	Tyrrell	Critical Ratio
Standard	7.16 (6.89)	9.41 (7.99)	3.99**
Var. 1	3.21 (3.45)	3.49 (3.48)	1.2
Var. 2	7.69 (7.46)	8.86 (7.8)	3.19**
Var. 3	8.38 (7.94)	13.52 (9.19)	8.87**

Also, Edmund's animal failed to exhibit even the slight dependency on food availability that Tyrrell had shown, with absolutely no significant correlations other than between genetic fitness and lifespan. This is further evidence that Edmund's eating habits may be its key disadvantage in "normal" worlds.

Variable	gen. fit.	water	cereal	cover	fruit	mates	prey	lifespan
gen. fit.	1.000							
water	0.033	1.000						
cereal	0.027	0.017	1.000					
cover	0.009	-0.027	-0.040	1.000				
fruit	0.032	0.016	-0.001	-0.053	1.000			
mates	0.049	0.032	-0.037	0.005	-0.052	1.000		
prey	0.014	0.057	-0.009	-0.004	-0.017	0.028	1.000	
lifespan	0.924**	0.019	0.052	0.013	0.027	0.015	0.030	1.000

Presumably, utilising this information, Edmund could be made to do still better against the ERP. However, the present comparison is sufficient to illustrate the use of quantitative methods as explicated in this paper, and to lead to several points of discussion on these methods.

6 Discussion

The previous section illustrates the utility of applying the quantitative methods of experimental psychology in several ways:

- as a means of comparison between two systems,
- as a means of indicating developmental progress,
- as a means of analysis of both domains and behaviour, and finally,
- as a means of communication to the reader.

We have illustrated also that these methodologies complement other, conventional forms of analysis and comparison. Where useful, qualitative analysis of the programs was carried out through observation of their behaviour, for example when we noticed the animal wasn't consistently sleeping in its den at night. Also, we used documentation of Tyrrell's intentions and experiences from his literature, and his actual strategies from his code. All of these elements combined to give us a fuller understanding of the strengths and weaknesses of both ERP and Edmund in this task domain.

The crux of these strengths and weaknesses is *design.* Tyrrell states in his thesis that he gave no consideration to learning based action selection mechanisms, because they could not survive in his environment (which he strove to make ecologically plausible) long enough to develop. Certainly animal behaviour, though universally adaptive to at least some extent, is highly constrained and biased. Gallistel, Brown, Carey, Gelman & Keil (1991) argue that ontogenetic learning in nature is reserved for information that cannot be acquired on an evolutionary time scale, and that what learning does exist is generally extremely specialised[7]. Certainly there is no such thing as an action selection mechanism that does not rely on design; any learned action selection still requires the developer selects the primitives and domain. Learning without bias is impossible (see (Kaelbling 1997) for a recent perspective.) The question is then, to what *extent* is action selection designed, and the issue of an agent architecture is not only whether it is capable of arbitrating behaviour, but to what extent it facilitates the design of that arbitration (Bryson in press).

As can be seen in the intermediate results of our own development, the setting of parameters on behaviour, even once it is defined plays a sizeable role in the performance of the system. For Tyrrell, new means of combination for weights and parameters was one of his major contributions in extending the Rosenblatt & Payton model, and is also listed as a major area for future work. Tyrrell had hoped to automate the selection of parameters via genetic algorithms, but found the task impossible. He determined the genetic approach failed in his situation for the following reasons:

- There is a very low signal to noise ratio (reflected in high variance) in the genetic fitness of any particular model.
- There were too many parameters, so the combinatorics were prohibitive.
- The optimal values of the parameters were too interdependent, which again slows the process of genetic learning.

Consequently, all of Tyrrell's parameters were set by hand, a particularly laborious approach since there were parameters connected with every lined action, perception and intermediate node. In the developing the Edmund animal, we severely reduced the number of parameters considered partly by simplifying the design, and partly by testing the extent the performance varied with respect to these values. Although we did not demonstrate it in this work, this sort of quantitative analysis could be used to select a few relevant factors for either off-line or ontogenetic learning. Thus a learning space that is biased for both fertility and relatively safety could be designed.

One drawback of applying experimental methods to robotics is implicit in the fact we chose to illustrate our techniques in a simulated environment. It is difficult to replicate experiments a sufficient number of times to establish statistically significant results on a real robot. This is particularly a problem if the operational criteria that has been selected is something that takes a long time to establish and is

[7]Though taking this notion to the extreme overlooks the lessons of the Baldwin effect, that individual plasticity facilitates evolutionary processes of change. See for example (Hinton & Nowlan 1987).

highly variable. The genetic fitness criteria in the Tyrrell simulation is an extreme case of each of these. The value is not determined until the conclusion of an animal's life, which takes thousands of individual decisions and time steps, and the value is highly dependent on external circumstances that fluctuated broadly, such as the coincidence of predators or the number of mates that had not wandered out of the animal's territory. These issues are precisely why experimental psychologists favour highly controlled laboratory environments with simple, automatically registered success criteria, such as appropriate bar presses and touch screen touches. The use of such controlled environments is antithetical to much of the autonomous robotics literature (eg. (Brooks 1991*b*)). However, we suggest it has a place as part of the description and evaluation of robotic systems, since it allows for precise description through the techniques we have outlined.

7 Conclusion

As mentioned in Section 2.2, many researchers believe animal intelligence has made its major advances after undergoing major reorganisation as a consequence of significant change in habitat. It is reasonable to expect that similar reorganisations will prove necessary to improve mobile robot design. In mobile robotics the major reorganisations described by Livesey (1986) correspond to the rewriting (or relearning) of an agent architecture. Although these may occur quite frequently, the roboticist does not have the luxury of the large number of individuals and evolutionary time scale that animals require to successfully reorganise and advance, given only general behavioural information. Consequently, the roboticist needs more information so that she can accelerate the development process.

In this paper we have argued that the quantitative methods of experimental psychology are one such source of information. We have described several techniques in detail: hypothesis testing, correlation, and regression; and also the design of experiments that best exploit them. We then proceeded to illustrate the use of these techniques to analyse a problem domain, develop an action selection mechanism to operate within it, and then evaluate that mechanism against others. Our conclusion is that these techniques are a rich source of information that can help speed the advance of mobile robotics, both in individual development and as a field.

References

Agre, P. E. & Chapman, D. (1988), What are Plans For?, AI memo 1050, MIT, Cambridge, MA.

Beer, R. D. (1989), Intelligence as Adaptive Behavior: An Experiment in Computational Neuroethology, Technical Report Technical Report 89-118, Center for Automation and Intelligent Systems Research, Case Western Reserve University.

Blumberg, B. M. (1996), Old Tricks, New Dogs: Ethology and Interactive Creatures, PhD thesis, MIT. Media Laboratory, Learning and Common Sense Section.

Brooks, R. A. (1991*a*), Intelligence Without Reason, *in* 'Proceedings of the Twelfth International Joint Conference on Artificial Intelligence', Morgan Kaufmann Publishers, Inc., Sydney, New South Wales, Australia, pp. 569–595.

Brooks, R. A. (1991*b*), 'Intelligence without Representation', *Artificial Intelligence* **47**, 139–159.

Brooks, R. A. & Stein, L. A. (1993), Building Brains for Bodies, Memo 1439, Massachusetts Institute of TechnologyArtificial Intelligence Lab, Cambridge, MA.

Bryson, J. (1995), The Reactive Accompanist: Adaptation and Behavior Decomposition in a Music System, *in* L. Steels, ed., 'The Biology and Technology of Intelligent Autonomous Agents', Springer-Verlag.

Bryson, J. (in press), Agent Architecture as Object Oriented Design, *in* M. P. Singh, ed., 'The Fourth International Workshop on Agent Theories, Architectures, and Languages (ATAL97)', Springer-Verlag.

Chapman, D. (1989), 'Penguins Can Make Cake', *AI Magazine* **10**(4), 51–60.

Cliff, D. (1992), Neural Networks for Visual Tracking in an Artificial Fly, *in* F. J. Varela & P. Bourgine, eds, 'Toward a Practice of Autonomous Systems: Proceedings of the First Conference on Artificial Life (ECAL91)', MIT Press, Cambridge, Massachusetts, pp. 78–87.

Collett, T. S. (1992), 'Landmark learning and guidance in insects', *Philosophical Transactions of the Royal Society of London, B* **337**, 295–303.

Connell, J. (1990), *Minimalist Mobile Robotics: A Colony-style Architecture for a Mobile Robot*, Academic Press, Cambridge, Massachusetts. also MIT TR-1151.

Coolican, H. (1994), *Research Methods and Statistics in Psychology*, second edn, Hodder & Stoughton.

Correia, L. & Steiger-Garção, A. (1995), A Useful Autonomous Vehicle With a Hierarchical Behavior Control, *in* F. Moran, A. Moreno, J. Merelo & P. Chacon, eds, 'Advances in Artificial Life (Third European Conference on Artificial Life)', Springer-Verlag, Berlin, Germany, pp. 625–639.

Devore, J. L. (1991), *Probability and Statistics for the Engineering Sciences*, Brooks/Cole Publishing, Pacific Grove CA.

Firby, J. (1987), An investigation into reactive planning in complex domains, *in* 'Proceedings of the National Conference on Artificial Intelligence (AAAI)', pp. 202–207.

Fontan, M. S. & Mataric, M. J. (1996), The Role of Critical Mass in Multi-Robot Adaptive Task Division, Technical Report CS-95-187, Brandeis University Computer Science. submitted to IEEE Transactions on Robotics and Automation.

Gallistel, C., Brown, A. L., Carey, S., Gelman, R. & Keil, F. C. (1991), Lessons From Animal Learning for the Study of Cognitive Development, *in* S. Carey & R. Gelman, eds, 'The Epigenesis of Mind', Lawrence Erlbaum, Hillsdale, NJ.

Gat, E. (1991), Reliable Goal-Directed Reactive Control of Autonomous Mobile Robots, PhD thesis, Virginia Polytechnic Institute and State University.

Georgeff, M. P. & Lansky, A. L. (1987), Reactive reasoning and planning, *in* 'Proceedings of the Sixth National Conference on Artificial Intelligence (AAAI-87)', Seattle, WA, pp. 677–682.

Harlow, H. F. (1958), 'The Nature of Love', *American Psychologist*.

Hayes, N. (1994), *Principles of Comparative Psychology*, Lawrence Erlbaum Associates.

Hendriks-Jansen, H. (1996), *Catching Ourselves in the Act: Situated Activity, Interactive Emergence, Evolution, and Human Thought*, MIT Press, Cambridge, MA.

Hinde, R. A. (1970), *Animal behaviour : a synthesis of ethology and comparative psychology*, McGraw-Hill.

Hinde, R. A. (1982), *Ethology*, Fontana Press.

Hinton, G. E. & Nowlan, S. J. (1987), 'How Learning Can Guide Evolution', *Complex Systems* **1**, 495–502.

Horswill, I. D. (1993), Specialization of Perceptual Processes, PhD thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Horswill, I. D. (1995), Visual routines and visual search, *in* 'Proceedings of the 14th International Joint Conference on Artificial Intelligence', Montreal.

Humphrys, M. (1996), Action Selection Methods Using Reinforcement Learning, *in* P. Maes, M. J. Matarić, J.-A. Meyer, J. Pollack & S. W. Wilson, eds, 'From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior', MIT Press, Cambridge, MA.

Jakobi, N. (forthcoming), 'Evolutionary robotics and the radical envelope of noise hypothesis', *Journal Of Adaptive Behaviour*.

Kaelbling, L. P. (1997), Why Robbie Can't Learn: The Difficulty of Learning in Autonomous Agents, *in* 'Proceedings of the 15th International Joint Conference on Artificial Intelligence', Morgan Kaufmann, Nagoya. IJCAI Computers and Thought Award talk.

Kitano, H., Asada, M., Kuniyoshi, Y., Noda, I. & Osawa, E. (1997), RoboCup: The Robot World Cup Initiative, *in* 'Proceedings of The First International Conference on Autonomous Agent', The ACM Press.

Levison, L. (1996), Connecting Planning and Action via Object-Specific Reasoning, PhD thesis, University of Pennsylvania. School of Engineering and Applied Science, Computer and Information Science Department.

Livesey, P. J. (1986), *Evolutionary Processes*, Vol. 1 of *Learning and Emotion: A Biological Synthesis*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Lorenz, K. (1973), *Foundations of Ethology*, Springer-Verlag, New York.

Maes, P. (1991), A bottom-up mechanism for behavior selection in an artificial creature, *in* J.-A. Meyer & S. Wilson, eds, 'From Animals to Animats', MIT Press, Cambridge, MA, pp. 478–485.

Maes, P., Matarić, M. J., Meyer, J.-A., Pollack, J. & Wilson, S. W., eds (1996), *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, MIT Press, Cambridge, MA.

McGonigle, B. (1991), Incrementing Intelligent Systems by Design, *in* J.-A. Meyer & S. Wilson, eds, 'From Animals to Animats', MIT Press, Cambridge, MA, pp. 478–485.

McGonigle, B. & Chalmers, M. (forthcoming), *The Growth of Intelligence in Complex Systems*, MIT Press, Cambridge, MA.

Milner, B., Corkin, S. & Teuber, H. L. (1968), 'Further analysis of the hippocampal syndrome: 14-year followup study of H. M.', *Neuropsychologia* **6**, 215–234.

Nilsson, N. (1984), Shakey the Robot, Technical note 323, SRI International, Menlo Park, California.

Nilsson, N. (1994), 'Teleo-Reactive Programs for Agent Control', *Journal of Artificial Intelligence Research* **1**, 139–158.

Perlman, G. & Horan, F. L. (1986), 'Report on |STAT Release 5.1 Data Analysis Programs for UNIX and MSDOS', *Behavior Research Methods, Instruments, & Computers* **18**(2), 168–176.

Premack, D. & Premack, A. J. (1983), *The Mind of an Ape*, W.W. Norton and Company.

Rosenblatt, K. & Payton, D. (1989), A fine-grained alternative to the subsumption architecture for mobile robot control, *in* 'Proceedings of the IEEE/INNS International Joint Conference on Nueral Netowrks'.

Savage-Rumbaugh, S., McDonald, K., Sevcik, R. S., Hopkins, W. D. & Rubert, E. (1989), 'Spontaneous Symbol Acquisition and Communicative Use by Pygmy Chimpanzees (Pan paniscus)', *Journal of Experimental Psychology: General* **115**(3), 211–233.

Singh, M. P., ed. (in press), *The Fourth International Workshop on Agent Theories, Architectures, and Languages (ATAL97)*, Springer-Verlag.

Steels, L. (1994), Building Agents with Autonomous Behavior Systems, *in* L. Steels & R. Brooks, eds, 'The 'artificial life' route to 'artificial intelligence'. Building situated embodied agents.', Lawrence Erlbaum Associates, New Haven.

Tanji, J. (1996), Involvement of motor areas in the medial frontal cortex of primates in temporal sequencing of multiple movements, *in* R. Caminiti, K. Hoffmann, F. Lacquaniti & J. Altman, eds, 'Vision and Movement: Mechanisms in the Cerebral Cortex', Vol. 2, Human Frontier Science Program, Strasbourg, pp. 126–133.

Terrace, H. S. & McGonigle, B. (1994), 'Memory and representation of serial order by children, monkeys and pigeons', *Current Directions in Psychological Science* **3**(6), 180–185.

Trullier, O., Wiener, S., Berthoz, A. & Meyer, J. A. (in press), 'Biologically-based artificial navigation systems: Review and prospects.', *Progress in Neurobiology*.

Tyrrell, T. (1993), Computational Mechanisms for Action Selection, PhD thesis, University of Edinburgh. Centre for Cognitive Science.

Webb, B. (1994), Robotic Experiments in Cricket Phonotaxis., *in* D. Cliff, P. Husbands, J.-A. Meyer & S. W. Wilson, eds, 'From Animals to Animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior', MIT Press, Cambridge, MA.

Whitehead, S. D. (1992), Reinforcement Learning for the Adaptive Control of Perception and Action, Technical Report 406, University of Rochester Computer Science, Rochester, NY.

Wilson, M. & McNaughton, B. (1994), 'Reactivation of Hippocampal Ensemble Memories During Sleep', *Science* **261**, 1227–1232.

Wooldridge, M. & Jennings, N. R. (1995), 'Intelligent Agents: Theory and Practice', *Knowledge Engineering Review*.