Chapter 3

AGENT LEARNING REPRESENTATION: ADVICE ON MODELLING ECONOMIC LEARNING

THOMAS BRENNER[1]

Max Planck Institute for Research into Economic Systems, Jena, Germany

Contents

[1] I want to thank John Duffy, Robert Marks, Alex Smajgl, Leigh Tesfatsion and an anonymous referee for helpful comments on earlier drafts of this chapter. The usual disclaimer applies.

Abstract

This paper presents an overview on the existing learning models in economic literature. Furthermore, it discusses the choice of models which should be used under various circumstances and how adequate learning models can be chosen in simulation approaches. It gives advice for using the many existing models and selecting the appropriate model for each application.

Keywords

economic learning, modelling

JEL classification: C63, D83

1. Introduction

In the last 20 years the variety of learning models used in economics has increased tremendously. This chapter provides an overview on these learning models. Furthermore, it classifies learning processes and gives tips on choosing amongst the various models.
There are many different ways in which such an overview can be presented and structured. The structure chosen here reflects two considerations: First, the main aim of this chapter is to help agent-based computational economists to choose the adequate learning model in their simulations. In giving such advice, we assume that agent-based computational economists intend to model human behaviour as realistically as possible. Other arguments in the context of choosing learning models are discussed in Section 5.2. However, the question of how real learning processes can be accurately modelled is the central concern of this chapter. As a consequence, the chapter is strongly based on research in psychology because psychologists have established most of the actual knowledge about human learning. Experimental economics has made increasingly larger contributions to this knowledge in recent years (a comprehensive overview is given in Duffy 2005). Nevertheless, most current knowledge stems from psychology.
Second, most researchers agree that there is no single universal learning model. Different learning processes take place in different situations (see experimental evidence presented in Duffy 2005). Thus, different learning models have to exist. In order to support agent-based computational economists in choosing a model, the learning situations have to be categorised and separate advice has to be given for each category. Many different categorisations are possible. The categorisation used here was developed earlier on the basis of the psychological literature (see Brenner 1999). It is based on the assumption that there is a hard-wired learning process that is common among all animals and a flexible learning process that requires features of the human brain. Other categorisations are possible and the specific choice in this overview is motivated and discussed in detail in Section 1.1.
The chapter proceeds as follows: The remaining introductory section gives a short historical overview and presents and discusses the categorisation of learning models used here. Subsequently, for each learning process class, the various models available are presented in Section 2 (non-conscious learning), Section 3 (routine-based learning), and Section 4 (belief learning). Section 5 discusses the possibility of a general learning model and addresses some basic issues in modelling learning, such as the complexity and validity of learning models, the distinction between individual and population learning, and the calibration of learning models. Section 6 concludes by giving detailed advice on how to adequately model the various ways of learning.

1.1 History of modelling learning

This short trip through history will focuses on mathematical learning models used in economics. Nevertheless, it is necessary to start with a short overview of psychological research on learning as the study of learning processes is mainly handled within psychology and many models that are used in economics are based on psychological findings. Furthermore, it was psychologists who developed the first mathematical models of learning.

1.1.1 Psychological research on learning

Psychologists started to study learning processes extensively approximately 100 years ago. At that time, psychology was dominated by the view that processes within the brain cannot be studied and that explanations of behaviour should be based purely on observable variables. Subsequently, psychologists identified two major learning processes: classic conditioning and operant conditioning. So far, classic conditioning has had little impact on economic discussion (an exception can be found in Witt 2001), although it is still extensively studied in psychology (an overview is given in Mackintosh 2003). It describes the development of new stimuli and reinforcers on the basis of existing ones and can, therefore, explain change in preferences (see Witt 2001). Mathematical models for this learning process have so far only been developed within psychology (see Rescorla & Wagner 1972), while economic literature has focused more on the process of operant conditioning. Most of the empirical studies in psychology on operant conditioning are conducted with animals. A general result that has been found is that actions that lead to rewards occur with a higher frequency in the future, while actions that cause punishment become less frequent. This kind of learning process is nowadays referred to as 'reinforcement learning' in economics. The first mathematical model of this learning process was developed within psychology by Bush and Mosteller (1955).

In the 1950s psychologists started a new line of research into learning processes. They studied the impact of social interaction and observation on learning. The basic argument was that people do not only learn from their own experience but also from the experience of others, meaning that the concept of reinforcement learning was transferred to interactions and observation. However, it was not only assumed that experience was exchanged between individuals but it was also claimed that people are able to understand the similarities and differences between other's and their own situation. Psychologists at that time entered the sphere of cognition and the resulting theory was called a social-cognitive learning theory (with the most prominent work by Bandura 1977).

Finally, in the last 20 years psychologists have concentrated on the processes of cognitive learning. Cognitive learning, in general, means the development of an understanding of real world processes and interrelations including the development of notions and expressions. Nowadays much research is done on the development of cognitions in children, such as the learning of languages, and logical thinking. However, formulations of the learning processes in the form of equations are absent as the processes are usually described using graphs, accompanied by verbal arguments or logical elements. The main topics centre on the development of structures and the integration of knowledge in the brain. Hence, this research is far removed from the type of decision making usually studied in economics with reference to learning processes.

A recent development is the use of neuro-science in the study of cognitive learning processes (see, e.g., Rumiati & Bekkering 2003). This research offers new information on the speed of information processing; the interaction between different stimuli that appear simultaneously; the extent to which different parts of the brain are involved in the processing of stimuli and similar aspects of learning. Again this research is, up to this point, of little use for modelling learning processes in economics. However, this may well change in the future.

1.1.2 Learning and optimisation

For a long time learning was a minor issue in economics. When economists started to show some interest in learning, they were mainly concerned with two issues. First, they established a normative learning model that described the optimal learning process, entitled Bayesian learning (see, e.g., Easley & Kiefer 1988 and Jordan 1991). Second, they developed models of learning in which behaviour converges towards the optimal behaviour in equilibrium. For quite some time, most economists who studied learning processes were mainly concerned with proving that learning converges towards the optimal behaviour. The first approach of this kind appeared in 1951 (see Brown 1951). After the proposal of the Nash equilibrium (Nash 1950) the question arose of how people come to play according to this equilibrium. Brown established a learning model called fictitious play for which Robinson (1951) could show that it converges to Nash equilibrium behaviour (later it was proved that this only holds under certain conditions; see Shapley 1964).

Many authors who model learning processes still want to show that learning processes converge towards optimisation (examples can be found in Bray 1982, Yin & Zhu 1990, Jordan 1991, Börgers & Sarin 1997, Dawid 1997 and Sarin & Vahid 1999). Often it is even argued that learning models can only be adequate if they converge, at least in the long run, towards optimising behaviour in a stationary situation. However, mainly caused by enormous experimental evidence, this claim is slowly disappearing from the debate. There are increasingly more works that study when and how behaviour predicted by learning models differs from optimal behaviour (see, e.g., Herrnstein & Prelec 1991, Brenner 1997, Brenner 2001, and Brenner & Vriend 2005).

Nevertheless, economists who model learning processes are still very much divided into two camps: those who prefer learning models that converge towards optimal behaviour and those who are not interested in optimality. In contrast to this, we argue here that finding out under which conditions the various existing models work best is more important (see Börgers 1996 for a similar argument).

1.1.3 Increasing variety of learning models

In the past few years, there has been a tremendous increase in the number of learning models used in economics. After experimental studies have repeatedly shown that the original economic learning models have been rejected in some experiments (see, e.g., Feltovich 2000), many economists, who modelled learning in economic contexts, developed their own model or their own variation of an existing model. Most of these models are based on introspection, common sense, artificial intelligence approaches or psychological findings. Nearly all of them are in some way or another set up *ad-hoc* without clear scientific justification.

In the meantime some approaches have tried to compare the suitability of different models on the basis of experimental data (see, e.g., and Feltovich 2000, Tang 2003 and Artifovic & Ledyard 2004). This topic is addressed and extensively discussed by Duffy (2005).

Independent of this discussion a few learning models have become dominant in economics, while others have been mainly neglected. The most prominent models are Bayesian learning, least-squares learning, as well as the learning direction theory, reinforcement learning, evolutionary algorithms, genetic programming, fictitious play and the learning model by

Camerer and Ho. There are different reasons for the dominance of these models, which range from being well-supported by empirical and experimental evidence to converging to optimal behaviour or reducing complexity. We present here more than these prominent models, although it is impossible to present all existing models and modifications of models.

1.2 Classification of learning models

A classification is always as beneficial as it is helpful for practical tasks. We consider here an agent-based computational economist who aims to explain features and dynamics of the economy on the basis of interaction between economic agents. For such an endeavour it is important to know the way in which economic agents behave and the adequate ways to model this behaviour in simulations. The task is to choose a learning model for a planned simulation study. Given the above assumption that the aim is to find the most realistic model, we have to ask what is the right learning model in a given situation.
However, in economic literature on learning this is not the only aim, and not even the most frequent one. Other aims are discussed in Section 5.2.1. Searching for realistic learning models, most information about real learning processes can be found in the psychological literature, thus building the basis for the classification proposed here.

1.2.1 Potential alternative classifications

Alternative ways to classify learning models should not be ignored here. There are at least three other options. First, one might classify learning models according to their origin. This would allow us to distinguish between psychology-based models, rationality-based models, adaptive models, belief learning models, and models inspired by computer science and biology. A classification of all learning models that are discussed here according to such a classification and according to the classification developed below is given in Table 1. Such a distinction informs the reader about the various sources of learning models, however, it does not help in choosing a model for simulations.
Second, we might classify learning models according to the economic fields in which they are usually applied. For example, macro-economists mainly use Bayesian learning and least-squares learning while reinforcement learning, fictitious play and learning direction theory are prominent among experimental economists. Meanwhile, evolutionary algorithms and genetic programming are frequently used in agent-based computational economics and game theorists seem to prefer fictitious play, replicator dynamics and other adaptive learning models. However, it is unclear why economists in different fields use different learning models. Obviously, economists who use mathematical analysis are restricted in their choice by the requirement of treatable models. The other differences seem to be historical in nature and it could be rather un-productive to support such differences by using them for a classification of learning models.
Third, one might look for existing classifications of learning models in economic literature. However, no classification is available that contains as many different learning models as discussed here. Usually only a few subjectively selected learning models are presented and discussed (see, e.g., Fudenberg & Levine 1998).

1.2.2 Proposed classification

The classification chosen here is based on the aim to assign realistic learning models to various situations. It is strongly based on psychological knowledge about learning.

It is not clear whether there is a fundamental mechanism within the brain that explains all learning methods. However, since neuro-physiologists and cognitive psychologists have not yet detected such a fundamental mechanism, learning methods can only be developed through empirical observations. Furthermore, it may be technically advantageous not to base all learning processes on one fundamental mechanism. Often it is simpler to describe learning processes on the basis of resulting changes in behaviour than to describe the probably complicated interaction of cognitive processes.

Thus, we are looking for information that helps us decide which learning model is suitable under certain conditions. We ignore the alternative option to search for the best learning model that describes all learning processes and is suitable under every condition. No such model exists, per se, and it is doubtful that there will ever be one.

While looking for a match between adequate learning models and situational characteristics, we are less interested in whether a model has structural attributes facilitating its use in a given situation. We are more interested in whether a learning model describes the relevant processes that occur in reality. Hence, we have to find out if various kinds of learning processes with different features exist in reality and how they occur.

1.2.3 Two ways of learning

Although the psychological literature on learning distinguishes (for historical reasons) between three kinds of learning processes, there are only two fundamentally different ways of learning. First, humans share with other animals a simple way of learning, which is usually called reinforcement learning. This kind of learning seems to be biologically fixed. If an action leads to a negative outcome – a punishment – this action will be avoided in the future. If an action leads to a positive outcome – a reward – it will reoccur. This kind of learning process has been extensively studied in psychology around 100 years ago with different kinds of animals (extensive literature can be found in Thorndike 1932 and Skinner 1938). It does not involve any conscious reflection on the situation. Hence, people are not always aware that they are learning.

In addition to reinforcement learning, people are able to reflect on their actions and consequences. We are able to understand the mechanisms that govern our surrounding and life; and we are able to give names to objects and establish causal relations that describe their interaction and nature. Nowadays, this is mainly studied in psychology under the label of learning and is referred to as cognitive learning.

These two kinds of learning are completely different. We argue – without having any empirical proof for this – that reinforcement learning is a mechanism that works in an automatically and continuous fashion. Subsequently, whatever we do is, instantaneously, guided by reinforcement learning. It seems likely that humans are endowed with the same basic mechanisms as animals and therefore learn according to the same hard-wired principles of reinforcement learning.

However, we are able to reflect on our actions and their consequences. This requires active thinking and, therefore, cognitive resources, which are scarce. Hence, we are not able to reflect on all our actions. Imagine if we would have to consider each move of each single muscle. We would not be able to live our life as we do. However, if we think about an action, we are able to overrule the law of reinforcement learning. We argue that the effect of cognitive learning on behaviour is stronger than the effect of reinforcement learning. But, in addition, we argue that we do not have the cognitive capacity to reflect on all our actions and therefore many actions are conducted on the basis of reinforcement learning.

1.2.4 Further distinction of learning processes

While reinforcement learning (or conditioning as it was originally named by psychologists) is well studied and understood, conscious learning processes are more difficult to grasp. Although various learning models exist in psychology, detailed knowledge on the formation of beliefs in the brain are missing.

Hence, there is some temptation to ignore the exact working of the brain and model some basic mechanisms of learning that are well established from empirical and experimental observations. Such models take a mechanistic perspective on learning. People are assumed to learn according to fixed mechanisms or routines. Therefore, we call the learning processes described by these models routine-based learning.

An example is the rule to imitate local people whilst in a foreign country for the first time. In this way one quickly learns about the traditions there and adapts behaviour. However, conscious learning is more than simply imitating the behaviour of others. Conscious learning usually means that we understand why this behaviour is advantageous, maybe how it developed and what are its suitable circumstances. This means that we associate meaning to our observations and build beliefs about relationships and future events. In order to distinguish these processes from simplified routine-based learning, it is defined as associative learning (in accordance with Brenner 1999) or belief learning in accordance with the term used in the economics literature.

All conscious learning is belief learning because, whenever people reflect upon their situation and learn about appropriate actions, they assign meanings to the gathered information, thus developing beliefs about relationships and future events. Routine-based learning is a simplification of real learning processes that makes life easier for the researcher and is applicable in a number of situations. The correct way would be to model the belief learning process, although the correct way is not always the appropriate way. This holds especially because we have little information about how belief learning processes should be modelled. Nevertheless, while using routine-based learning models, we should keep in mind that they only represent approximations.

2. Modelling non-conscious learning

According to the categorisation proposed here, all learning processes that occur without individuals being aware of them are labelled non-conscious learning. In psychology two such learning processes are identified: classical conditioning and operant conditioning (also called reinforcement learning). As mentioned above, the discussion is restricted here to the

process of reinforcement learning as classical conditioning is rarely addressed in economic literature (a discussion of modelling classical conditioning can be found in Brenner 1999, Ch. 3 and 5). However, it has to be mentioned that if we equate non-conscious learning with reinforcement learning, we depart from the traditional psychological notion of reinforcement learning. In psychology, reinforcement learning was established at a time in which behaviourism was dominant, which meant that models which explicitly considered the internal functioning of the cognitive processes were argued by psychologists to be pure speculation and therefore should be avoided. Hence, they developed models of learning processes that saw decisions as being outcomes of visible processes, i.e. stimulus-response relations. However, these models have not excluded the possibility that there might be cognitive processes in the background that cause visible changes in behaviour. They only hold that these processes should not be explicitly included in the models.
Non-conscious learning, as it is defined here, applies only to those learning processes in which no cognitive reflection takes place. The analogy that we draw here comes from the fact that most psychological studies of conditioning have been based on animal experiments which claim that animals mainly learn non-cognitively. Hence, we argue that reinforcement learning models should be suitable for modelling non-cognitive learning processes in humans. In psychology it is frequently argued that individuals learn according to reinforcement learning if they do not reflect on the situation (see, e.g., Biel & Dahlstrand 1997).
Reinforcement learning is based on an initial frequency distribution among various possible actions. The origin of this frequency distribution has to be explained by other means, as it has been mainly neglected in the literature. Reinforcement learning means that actions are chosen randomly according to the current frequency distribution. If an action leads to a reward (positive outcome) the frequency of this action in future behaviour is increased. If an action leads to a punishment (negative outcome) the frequency of this action is decreased.

2.1 Existing models

In economics three frequently used models that describe reinforcement learning are the Bush-Mosteller model, the principle of melioration, and the Roth-Erev model. All three models best capture the major characteristic of reinforcement learning: the increase in the frequency of behaviours that lead to relatively better results and the slow disappearance of behaviours if reinforcement is removed. All these models are inspired by psychological research on reinforcement learning. However, as discussed above, non-conscious learning is not the same as reinforcement learning. Therefore, we have to depart from the psychological literature here and assess whether these models actually describe a non-conscious learning process.
The three models do differ in their details. Melioration learning assumes that the learning process is based on the average experience of each behaviour in the past. The Bush-Mosteller model and the Roth-Erev model assume that the change of behaviour at each point in time is determined by the current outcome together with the previously determined frequency distribution, which are used to determine an updated frequency distribution

for current action choice. Hence, while the Bush-Mosteller model and the Roth-Erev model only require the individual to store the actual frequencies of the possible actions, melioration learning requires them to also remember past events. Furthermore, it requires individuals to calculate averages. Herrnstein developed the melioration principle in the light of experimental observations (see Herrnstein 1970 and Herrnstein & Prelec 1991). However, given their laboratory settings, such behaviour is probably conscious. Hence, the melioration concept seems to fit better into the modelling of routine-based learning and will be further discussed there. It does not seem to be adequate to model non-conscious learning processes.

Juxtaposing the Bush-Mosteller model and the Roth-Erev model, one may observe that they have the same fundamental structure. The Bush-Mosteller model was set up in 1955 by psychologists according to the psychological knowledge on operant conditioning (see Bush & Mosteller 1955). It was adapted to economics by Cross about 20 years later (see Cross 1973 and 1983). Arthur (1991) generalised the model by allowing for different developments of the learning speed during the learning process. He called his learning model 'parameterized learning automaton'. The two extreme cases that are included in Arthur's model are a constant learning speed and a hyperbolically decreasing learning speed. The former border case is identical to the Bush-Mosteller model. The latter border case later became the original Roth-Erev model. Arthur (1991) developed a very flexible model and discussed the meaning of different learning speeds. However, all these developments did not catch much attention within economics and it is to the merit of Roth and Erev to have reestablished reinforcement learning in economics.

The major difference between the Bush-Mosteller model and the original Roth-Erev model is the speed of learning, which was already discussed by Arthur (1991). In the Bush-Mosteller model the speed of learning remains constant. This means that an individual with a lot of experience in a situation reacts in the same way to a new experience as an individual with no former experience. The original Roth-Erev model assumes that the learning speed converges hyperbolically to zero while experience is collected. Psychological studies describe the aspect of spontaneous recovery (see Thorndike 1932), which means those actions that have been abandoned because of unpleasant results are quickly taken into the individual's behavioural repertoire again if the individual experiences positive outcomes resulting from these actions. This spontaneous recovery is captured by the Bush-Mosteller model but not by the original Roth-Erev model. However, Roth and Erev (1995) have modified their original model by including the aspect of forgetting, so that it also captures the process of spontaneous recovery.

A second difference is that the Bush-Mosteller model can also deal with negative payoffs, while the Roth-Erev model and the parameterized learning automaton are only able to use positive payoffs. The original psychological studies show that reinforcement learning has different characteristics for positive (rewarding) and negative (punishing) outcomes. The Bush-Mosteller model is able to capture these effects and accordingly leads to different predictions (see Brenner 1997). Thus, it is the only learning model that does not contain cognitive elements and is able to reproduce all features of reinforcement learning that have been identified in psychological studies.

As the parameterized learning automaton and the Roth-Erev model are described by Duffy

(2005), only the Bush-Mosteller model is described here in detail.

2.2 Bush-Mosteller model

At the beginning of the last century reinforcement learning became a central topic in psychology (cf. the previous section). This eventually led to the development of a mathematical learning model by Bush and Mosteller (1955). Their model is based on the considerations of Estes (1950) who took the first steps towards a mathematical formulation of reinforcement learning. It is based on the idea of representing behaviour by a frequency distribution of behaviour patterns given by a probability vector $\mathbf{p}(t)$ $(= (p(a,t))_{a \in \mathcal{A}})$. This vector assigns a probability $p(a,t)$ $(0 \leq p(a,t) \leq 1$, $\sum_{a \in \mathcal{A}} p(a,t) = 1)$ to each behavioural alternative a $(a \in \mathcal{A})$ at each time t. The term $p(a,t)$ is sometimes called habit strength. The Bush-Mosteller model is a stochastic model that predicts probabilities for the occurrence of behaviour patterns rather than the behaviour pattern itself.

The probability vector $\mathbf{p}(t)$ changes during the learning process according to the theory of reinforcement. Bush and Mosteller distinguished only between rewarding and punishing outcomes, but not within both classes. Cross (1973) further developed the Bush-Mosteller model by answering the question of how to deal with rewards and punishments of different strength. He placed the models into an economic context and so defined the reinforcing character of an event by the utility to which it gives rise. In doing so, he assumed that the impact of an outcome is monotonously increasing in its utility. However, Cross also eliminated the punishing character of events, because in economics it is assumed that utilities can be linearly transformed and negative utilities values can be avoided without a loss of generality as long as they have a finite lower bound. He overlooked that reinforcement learning works in a different manner for those situations in which agents are exposed to punishing outcomes compared to those situations in which they are exposed to rewarding outcomes. Therefore, in reinforcement learning it matters whether punishing or rewarding outcomes motivate learning.

This shortcoming of Cross's version of the Bush-Mosteller model has been overcome by the work of Börgers and Sarin (1997) and Brenner (1997). Only this version of the Bush-Mosteller model, called the generalised Bush-Mosteller model here, is described here (a discussion of all versions can be found in Brenner 1999, Ch. 3). Reinforcement strengths are defined in such a way that all rewarding outcomes are reflected by positive reinforcement strengths, while all punishing outcomes are reflected by negative reinforcement strengths. Apart from this, the generalised Bush-Mosteller model is identical to the version proposed by Cross. The change in the probability $p(a,t)$ of the individual to realise action a is given by

$$p(a,t+1) = p(a,t) + \begin{cases} \nu(\Pi(t)) \cdot (1 - p(a,t)) & \text{if} \quad a = a(t) \\ -\nu(\Pi(t)) \cdot p(a,t) & \text{if} \quad a \neq a(t) \end{cases} \tag{1}$$

if action $a(t)$ is realised and the resulting reinforcement strength $\Pi(t)$ is positive, and by

$$p(a,t+1) = p(a,t) + \begin{cases} -\nu(-\Pi(t)) \cdot p(a,t) & \text{if} \quad a = a(t) \\ \nu(-\Pi(t)) \cdot \frac{p(a,t)\, p(a(t),t)}{1-p(a(t),t)} & \text{if} \quad a \neq a(t) \end{cases} \tag{2}$$

if action $a(t)$ is realised and the resulting reinforcement strength $\Pi(t)$ is negative. $\nu(\Pi)$ is a monotonously increasing function in Π ($\Pi > 0$) with $\nu(0) = 0$ and $0 \leq \nu(\Pi) \leq 1$. A reinforcement strength of $\Pi = 0$ can be interpreted as the aspiration level (as done in Börgers & Sarin 1997).
Usually, a linear formulation $\nu(\Pi) = \nu \cdot \Pi$ is used, so that the learning process is described by

$$p(a, t+1) = p(a,t) + \begin{cases} \nu \cdot \Pi(t) \cdot (1 - p(a,t)) & \text{if} \quad a = a(t) \wedge \Pi(t) \geq 0 \\ \nu \cdot \Pi(t) \cdot p(a,t) & \text{if} \quad a = a(t) \wedge \Pi(t) < 0 \\ -\nu \cdot \Pi(t) \cdot p(a,t) & \text{if} \quad a \neq a(t) \wedge \Pi(t) \geq 0 \\ -\nu \cdot \Pi(t) \cdot \frac{p(a,t) \cdot p(a(t),t)}{1 - p(a(t),t)} & \text{if} \quad a \neq a(t) \wedge \Pi(t) < 0 \end{cases} . \tag{3}$$

All versions of the Bush-Mosteller model assume that an outcome has an impact on the frequency distribution $\mathbf{p}(t)$ in the moment of its occurrence only. This means that individuals do not remember previous actions and outcomes. The past is implicitly contained in the frequency distribution $\mathbf{p}(t)$. Learning is assumed to be a Markov process.
This model can only be applied to situations in which individuals have to choose repeatedly between a finite number of alternative behaviours, such as a set of different actions or a number of real-valued actions, e.g., prices. It cannot be applied to situations in which individuals have to choose a value from a set of infinite cardinality, such as an interval of possible prices. Choosing a real value within an interval implies conscious thinking, because the very notion of real value is a cognitive concept and must be consciously learned.

3. Modelling routine-based learning

It was extensively discussed above that routine-based learning models are approximations for the real conscious learning processes, as they are based on the identification of some simple fundamental principles of learning. These principles are deduced in economic literature either from experimental observations, from ad-hoc reasoning, or from some considerations on optimal learning. They might describe learning quite accurately under certain circumstances. However, they are never able to describe learning in each situation because people are capable of complex reasoning and of understanding the potentially complex environment that they face. Unfortunately, psychologists still have quite a vague comprehension of reasoning and understanding processes. Nevertheless, it is clear that these processes are not simple and that they involve the development of concepts and beliefs (see, e.g. Anderson 2000).
We define routine-based learning models as those models in which there is a direct connection from the agent's experiences and observations to their behaviour. All models that include beliefs and their development over time are seen as potential candidates for modelling, what is called here, associative or belief learning (they are discussed in the next section). We claim that there will never be a routine-based learning model that accurately describes the conscious learning process in all circumstances.

Nevertheless, under certain circumstances, routine-based learning models may be an adequate and simple description of learning. Several studies have shown that individuals tend to stick to their beliefs even if there is some evidence that falsify them (see, e.g., Luchins 1942 and Anderson 2000). Some strands of psychological research have shown that individuals apply simple rather than optimal routines in decision making (see, e.g., Gigerenzer & Selten 2001).

Although we have to be aware of the restrictions of routine-based learning models, it might be advantageous to search for a routine-based learning model in order to describe behaviour. In this case, the only way to guide this search is empirical and experimental evidence. The aim is not to have a detailed description of the learning process but to find a model that accurately represents the dynamics of these actions.

Various learning models have been put forward in the economics literature that fit into this category. The most frequently used models are represented here; some of them model one aspect of learning, while others combine a number of mechanisms within one model. Furthermore, there are models that describe learning on the individual level, while others describe them on a population level. These issues are discussed in Section 5.2. Here, the most prominent models are presented sequentially.

3.1 Experimentation

The standard form of learning by experimentation is the trial-and-error principle. However, it is not sufficiently specific to be called a model. It requires specification of whether all possible actions are tried, how often they are tried before they can be called an error and what an error means. Therefore, experimentation is usually included in other models as an additional factor. Nowadays, for almost all learning and decision models, some variants exist that include experimentation. Even utility maximisation has been expanded to include these elements, such as errors or individual differences in the evaluation of actions, which are similar to experimentation (see Brenner & Slembeck 2001 for a detailed presentation). The inclusion of experimentation in other models will be discussed during the presentation of these models below. Here two concepts will be presented that are only based on experimentation: the concept of $S(k)$-equilibria and the learning direction theory.

Experimentation is based on the argument that through choosing different actions individuals can collect information on the consequences of these actions. In the literature it is sometimes argued that there is a trade-off between experimentation and exploitation. This means that an individual can either experiment to obtain further information on the situation or exploit the information collected in the past using this as a basis to choose the best action.

A simple form of this kind of behaviour is the proposed behaviour by Osbourne and Rubinstein (1995). They argued that individuals choose each possible action k times. After this initial phase they choose the action that has led to the highest average payoff whenever they face the same situation. What remains unclear in this approach is the question of how k is determined. This model has also not been tested empirically or experimentally.

More evidence exists in favour of the learning direction theory (see, e.g., Berninghaus &

Ehrhart 1998 and Grosskopf 2003). The learning direction theory was proposed by Selten (see Selten & Stoecker 1986 and Selten 1997). Learning direction theory can only be applied if individuals are confined to choosing from a set of alternatives that can be ordered in a meaningful way, or if, at least, individuals are able to separate the alternatives that increase performance from those which decrease performance each time. Moreover, learning direction theory assumes that individuals are able to identify whether their last action was pitched too high or too low in this order of possible actions. Given these assumptions, learning direction theory states that individuals will change their behaviour in the direction in which they expect their own performance to increase, or stay with the same behaviour. Such a learning procedure has some similarities with gradient methods used in optimisation problems, although in the case of the learning direction theory there does not necessarily have to be something like a potential function. In other words, learning direction theory states that individuals change their behaviour only in a way that increases their payoff. As long as the situation is easy to understand, such a statement is straightforward. So it is no surprise that the theory has been confirmed in many experiments. However, the implications of learning direction theory are rather weak.

3.2 Melioration and experience collection

It is well-known that individuals memorise their experience with certain situations and use this experience to choose an adequate action if they face the situation again. Individuals even transfer experience between different situations that are perceived to be similar. In the psychological literature it is argued that probabilities and values are assigned to outcomes, both determined by previous experience.

In economics, various models have been developed that describe the collection of knowledge about a situation. Some of them are based on statistical considerations about how people should learn in an optimal way. The two most widespread models of this kind are Bayesian learning and least-squares learning. In contrast, some models are built according to how the modeller thinks people learn in reality. These include the models of myopic learning, fictitious play and melioration learning, which is based on experimental findings.

Except for melioration learning, an element that all these models have in common is that individuals not only learn about the results of their actions but also about the probabilities of the actions of other agents or events. Hence, although Bayesian learning, least-squares learning and fictitious play can be used to describe a learning process in which the outcomes of different actions are memorised, they are also able to describe a learning process in which beliefs and hypotheses about the situation are developed. They are suited to describe what is called belief learning here. Therefore, they are discussed in the next section. The only learning model that is only designed to describe only the memorising of experience with different actions is melioration learning. However, it should be stated here that we regard also fictitous play as a relevant model for experience collection.

Melioration learning, although it is never related to the model of fictitious play in the literature, is a special case of gradual convergence to fictitious play. Melioration learning was developed to represent reinforcement learning (see Herrnstein 1970, Vaughan & Herrnstein 1987 and Herrnstein & Prelec 1991). However, it is argued here that it is less adequate to model what was called non-conscious learning above. It seems to be more adequate to model the routine-based learning process of experience collection.

The dynamics of melioration learning have been formulated mathematically by Vaughan and Herrnstein (1987). For a case with two possible actions a and $\tilde{a}$, they describe it as a time continuous adjustment process of the form

$$\frac{dp(a,t)}{dt} = \nu\Big(\bar{u}(a,t) - \bar{u}(\tilde{a},t)\Big) \tag{4}$$

where $p(a,t)$ denotes the probability of income spent on activity a, $\bar{u}(a)$ denote the average utility from activity a in the past, and $\nu(\cdot)$ is a monotonously increasing function with $\nu(0) = 0$.

Vaughan and Herrnstein (1987) neglect cases in which the individuals can choose between more than two alternative actions. Nor do they define the average utilities $\bar{u}(a,t)$ in detail. A discussion of the average utility or payoff is presented by Brenner and Witt (2003). They define $\mathcal{T}_a(t)$ as the set of moments of time in which an individual has realised action a and also has a memory of it at time t. The term $k_a(t)$ denotes the number of these occasions. Consequently the average utility $\bar{u}(a,t)$ is given by

$$\bar{u}(a,t) = \frac{1}{k_a(t)} \cdot \sum_{\tau \in \mathcal{T}_a(t)} u(a,\tau) \ . \tag{5}$$

Brenner and Witt (2003) also claim that it is more adequate to multiply equation (4) by $p(a,t) \cdot (1 - p(a,t))$ instead of the artificial additional condition that the dynamics stop if $p(a,t)$ becomes smaller than zero for at least one action a. Consequently equation (4) can be written as

$$\frac{dp(a,t)}{dt} = p(a,t) \cdot \Big(1 - p(a,t)\Big) \cdot \nu\Big(\bar{u}(a,t) - \bar{u}(\tilde{a},t)\Big) \ . \tag{6}$$

This approach can be easily expanded to a situation in which the individual has more than two options to choose. Let us assume that there is a set $\mathcal{A}$ of alternative actions. Furthermore, let us assume that the function $\nu(\cdot)$ is linear. Then, the dynamics of melioration learning is given by

$$\frac{dp(a,t)}{dt} = p(a,t) \cdot \nu \cdot \Big(\bar{u}(a,t) - \sum_{\tilde{a} \in \mathcal{A}} p(\tilde{a},t) \cdot \bar{u}(\tilde{a},t)\Big) \ . \tag{7}$$

Equation (7) describes a replicator dynamic (compare to Equation (13)). At the same time, it also represents a special case of a gradual convergence to fictitious play. The utility or payoff $u(a,t)$ is calculated for each action a on the basis of a finite memory. However, the action that caused the highest payoffs in the past is not immediately chosen. Instead, the behaviour converges towards the choice of the action with the best experience. Through this process, individuals keep experimenting as long as one action does not supersede the others for a relatively long time. Such a model seems to be a good choice for modelling a realistic combination of experience collection and experimentation, but it is nevertheless quite simple. Alternatively, the average payoffs $\bar{u}(a,t)$ might also be calculated as exponentially weighted averages of the past experience according to

$$\bar{u}(a,t) = \frac{1-\beta}{1-\beta^{(t-1)}} \sum_{\tau=0}^{t-1} \beta^{(t-1-\tau)} \cdot u(\tau) \cdot \delta(a(\tau) = a) \tag{8}$$

where $u(\tau)$ is the utility obtained by the individual at time τ, $a(\tau)$ is the action taken at time τ, and $\delta(a(\tau) = a)$ is 1 if $a(\tau) = a$ and otherwise 0. β is a parameter that reflects the time-horizon of the memory.

3.3 Imitation

The process of imitation is often used to describe learning processes in economics. Yet no general model exists that describes imitation. Each author who considers imitation an important aspect makes her own assumptions about the process. Most models of imitation found in economic literature assume that the individuals are able to observe the actions of other individuals and their resulting outcomes. Furthermore, the individuals are assumed to use this information in order to take actions that lead to a better outcome. This contrasts with the recent psychological literature on imitation. There, imitation is seen as an innate process: Children are found to imitate behaviours that have no real advantage. Nevertheless they are also able to concentrate on the crucial features for success and neglect minor aspects as well as learn from unsuccessful behaviour (see Rumiati & Bekkering 2003 for a condensed overview). However, the psychological research has not produced learning models that could be used in economics.
Therefore, the different models that have been developed in economics are discussed below. These models differ in various characteristics:

- In some models it is claimed that a certain number of individuals are observed, where these are located next to the observer (see, e.g., Eshel, Samuelson & Shaked 1998) or they are randomly picked from the whole population (see, e.g., Duffy & Feltovich 1999 and Kirchkamp 2000). In other models the entire population is observed (see, e.g., Morales 2002). Because the situation determines how many individuals are observed in reality, the existence of various models is justified.

- Some models calculate the average utility of each action based on observations (see, e.g., Eshel, Samuelson & Shaked 1998). Other models claim that individuals imitate the one who has obtained the highest utility of those observed (see, e.g., Nowak & May 1993, Hegselmann 1996, Vega-Redondo 1997 and Kirchkamp 2000). Vega-Redondo adds noise to this assumption. Finally, in some models it is assumed that only one other individual is observed at any time and the payoff or utility obtained by this individual is compared to their own payoff. Then, either a stochastic approach is chosen whereby it is more likely to imitate those with a higher the difference between the other's and their own utility (see, e.g., Witt 1996) or the other individual's behaviour is imitated whenever it has given rise to a higher utility (see Schlag 1998 for a discussion of these different rules). There is no empirical or experimental study that examines which of these many models is more realistic.

Although the psychological literature offers no mathematical model, it does offer some conceptual help. Imitation learning is discussed in psychology under the label of observational learning (see Bandura 1977). It is discussed how the attention of people is drawn to the experience made by others and how this experience is transferred to one's own situation. This literature implies that we may treat imitation via the models of experience collection that have been described above. Furthermore, such a modelling would include the process of communication. The above Equation (8) can be modified such that it also contains the experience gathered by other individuals. Then, the only question that has to be answered is the question of how much of other individuals' experience is observed and considered in decision making. The answer is context-dependent and has to be found for each situation separately.

A very simple model that combines the routine-based processes of experimentation, experience collection and imitation/communication would consist of two processes: First, an exponentially weighted average of the past experience with each action a is built, including the experience of all N individuals in the population:

$$\bar{u}_i(a,t) = \frac{1-\beta}{1-\beta^{(t-1)}} \sum_{\tau=0}^{t-1} \left[\beta^{(t-1-\tau)} \cdot \sum_{j=1}^{N} \sigma(i,j) \cdot u_j(\tau) \cdot \delta(a_j(\tau) = a) \right] \tag{9}$$

where $u_j(\tau)$ is the utility obtained by individual j at time τ, $a_j(\tau)$ is the action taken by individual j at time τ, and $\sigma(i,j)$ is the weight with which individual i includes the experience of individual j in her own expectation about the future. These weights have to satisfy

$$\sum_{j=1}^{N} \sigma(i,j) = 1 \tag{10}$$

for each individual i. Besides this, the weights have to be determined specifically for each situation. The change of behaviour can then be modelled by Equation (7). The above can however also be used to learn about the circumstances, such as the behaviour of others or any rules of the situation, as it is modelled in fictitious play.

3.4 Satisficing

The concept of satisficing can be found in many learning models in the literature. It was first proposed by Simon (1957). Since then, many models have been proposed that describe learning on the basis of the satisficing principle (for a detailed description of the satisficing principle see Simon 1987). Many of these models, however, are based on one of the routine-based learning processes above and contain satisficing as an additional aspect. The satisficing principle is based on the assumption that individuals have an aspiration level. This means that they assign a value z to each situation for the payoff or utility that they expect to obtain. If the actual payoff or utility is above this value they are satisfied, while an outcome below this value dissatisfies them. In order to model satisficing three things have to be specified: 1) the aspiration level, 2) the dependence of behavioural changes on dissatisfaction, and 3) how the new action is chosen if the current one is abandoned because of dissatisfaction.

Aspiration levels have been studied intensively in psychology in the 1940s and 1950s (see, e.g., Festinger 1942 and Thibaut & Kelley 1959 and Lant 1992 and Stahl & Haruvy 2002 for economic studies). From these empirical studies we know that the aspiration level of an individual changes in time. On the one hand, it depends on the outcomes of the individual's own behaviour. These experiences are considered more seriously the more recent they are (see Thibaut and Kelley 1959). On the other hand, it depends on the performances of others (see Festinger 1942). Here again the influence of outcomes decreases with increasing time but also depends on the similarity of these people and their situation with one's own position.

In the literature, three kinds of aspiration levels can be found (see Bendor, Mookherjee & Ray 2001 for an overview). First, some authors assume for simplicity a constant aspiration level z (see, e.g., Day 1967 and Day & Tinney 1968). Second, some authors assume an aspiration level $z_i(t)$ of an individual i that adapts towards the payoffs currently obtained by this individual (see, e.g., Witt 1986, Mezias 1988, Gilboa & Schmeidler 1995, Pazgal 1997, Karandikar, Mookherjee, Ray & Vega-Redondo 1998, and Börgers & Sarin 2000). Third, some authors let the experiences of others influence the formation of individual aspirations, as it was proven in psychological studies. This additional impact is called 'social influences' (see, e.g., Dixon 2000 and Mezias, Chen & Murphy 2002).

The assumption of a constant aspiration level is rare in the recent literature. Therefore, we focus on the two other approaches here. The most common way of modelling (see, e.g., Karandikar, Mookherjee, Ray & Vega-Redondo 1998 and Börgers & Sarin 2000) an aspiration level that adapts to the personal experience is based on the equation

$$z(t+1) = \lambda \cdot z(t) + (1-\lambda) \cdot \pi(t) \ . \tag{11}$$

In this equation λ determines how much the new experience influences the aspiration level $z(t)$ and $\pi(t)$ is the payoff obtained by the individual at time t. Alternative models mainly replace the payoff $\pi(t)$ in Equation (11) by other variables such as the utility obtained at time t or the maximal or average payoff in the past.

The influence of the experience of other individuals can be included, for example (see Mezias, Chen & Murphy 2002), by introducing a social influence $\pi_{soc}(t)$:

$$z_i(t+1) = \lambda_1 \cdot z_i(t) + \lambda_2 \cdot [\pi_i(t) - z_i(t)] + \lambda_3 \cdot \pi_{soc}(t) \ , \tag{12}$$

where λ_1, λ_2 and λ_3 determine the strengths of the different influences on the aspiration level and have to add up to one: $\lambda_1 + \lambda_2 + \lambda_3 = 1$. The social influence $\pi_{soc}(t)$ can be defined as the average payoff that other individuals obtain at time t. However, it might also include further social aspects: An individual might put her own aspiration relatively higher or lower to what others reach. Furthermore, other individuals might vary in importance in the formation of the individual's own aspiration.

A further complication is that different assumptions about the reaction to dissatisfaction can be found in the literature. In general it is argued that the probability to change behaviour increases with the degree of dissatisfaction but never reaches one (see Palomino & Vega-Redondo 1999 and Dixon 2000). Usually a linear increase in the probability of changing behaviour is assumed (see, e.g., Börgers and Sarin 2000). However, other forms are possible as well.

Finally, a satisficing model has to specify how the new action is chosen if the current action is abandoned because of dissatisfaction. If there are only two alternative actions, this specification is straight-forward. In the case of more than two alternative actions there are two options. First, one of the other routine-based learning models can be used to determine the new choice. Through this, the satisficing principle can be combined with other concepts. Second, the new choice can be determined randomly.

3.5 Replicator dynamics and selection-mutation equation

The replicator dynamics (see Hofbauer & Sigmund 1984) is the basis for evolutionary game theory. It originates from biology and simply states that behaviours that are fitter than average occur more frequently and behaviours that are worse than the average occur less frequently. This is mathematically given by

$$\frac{ds(a,t)}{dt} = \nu(t) \cdot s(a,t)\Big[\Pi(a,t) - \langle\Pi(t)\rangle\Big] \ , \tag{13}$$

where $\nu(t)$ denotes the speed of the process, $\Pi(a,t)$ is the average outcome obtained by those individuals that show behaviour a at time t, and $\langle\Pi(t)\rangle = \sum_{a\in\mathcal{A}} s(a,t)\cdot\Pi(a,t)$ is the average outcome in the whole population at time t. The replicator dynamics describes the selection process in biological evolution. There, $\nu(t)$ is called the selection pressure, meaning the velocity of the elimination of less fit species. $\Pi(a,t)$ is the fitness of the species a at time t.

The selection-mutation equation (see Eigen 1971), also called Fisher-Eigen equation, also originates from biology. In addition to the selection process that is captured by the replicator dynamics, it also captures the mutation process. The selection-mutation equation can be written as (see Helbing 1995)

$$\begin{aligned} \tfrac{ds(a,t)}{dt} \;\; &= \;\; \textstyle\sum_{\tilde{a}\in\mathcal{A}} \Big[\omega(a|\tilde{a},t)\cdot s(\tilde{a},t) - \omega(\tilde{a}|a,t)\cdot s(a,t)\Big] \\ &\quad +\nu(t)\cdot s(a,t)\Big[\Pi(a,t) - \langle\Pi(t)\rangle\Big] \ . \end{aligned} \tag{14}$$

The first term on the right-hand side of Equation (14) represents the mutation processes. The mutation matrix $\omega(a|\tilde{a},t)$ defines the probability of a mutation from genetic variant $\tilde{a}$ to genetic variant a. The mutation matrix has to be chosen according to the biological probabilities of crossovers, mutations and other similar processes. In an economic context it can be chosen according to the probabilities of individuals randomly switching from one choice to another. The second term on the right-hand side of Equation (14) corresponds exactly to the relicator dynamics.

The replicator dynamics and selection-mutation equation are mainly used in mathematical analysis of learning processes because they are, in contrast to most other common learning models, analytically treatable. We are not aware of any experimental accuracy test, so little can be said whether they accurately represent learning processes. However, the selection-mutation equation is, at least, a flexible formulation. Defining the mutation matrix $\omega(a|\tilde{a},t)$ allows the inclusion of various aspects into the model. It is even possible to make this matrix, and thus the experimentation of individuals, dependent on the actual situation (such as the satisfaction of the individuals or their knowledge about potential improvements). However, these possibilities have so far not been examined closely in the literature.

3.6 Evolutionary algorithms

The evolutionary algorithms of Rechenberg and Holland (see Rechenberg 1973 and Holland 1975) are based on the same biological basis as the selection-mutation equation. However, Equation (14) is not able to exactly represent the dynamics of evolutionary algorithms because the selection process is modelled differently in evolutionary algorithms. Furthermore, evolutionary algorithms explicitly describe the development of each individual and its replacement in the next generation (a detailed description of genetic algorithms can be found in Duffy 2005; evolutionary strategies are thoroughly described in Beyer 2001). While the replicator dynamics and the selection-mutation equation dominate mathematical evolutionary game theory, it has become common to use evolutionary algorithms in economic simulations (see Holland & Miller 1991, Arifovic 1994 and Dawid 1996 for some path breaking works). A extensive representation of genetic algorithms and a discussion on the interpretation of these as learning processes can be found in Duffy 2005. Hence, we take up only two issues here that are rarely considered in computational economics but are important for using evolutionary algorithms to represent economic learning: learning aspects not represented by evolutionary algorithms and the difference between genetic algorithms and evolutionary strategies.

While in computational economics the analogy between genetic algorithms and real learning processes is widely accepted, this does not hold on a more general level where the similarities between biological evolution and cultural evolution, based on learning, are controversially discussed (see, e.g., Maynard Smith 1982, Hallpike 1986, Witt 1991, and Ramstad 1994). Some differences between biological evolution and learning processes also hold for using genetic algorithms in modelling economic learning. The main difference (for other differences see Brenner 1998) is that evolutionary algorithms contain a limited type of memory. Past experience is only remembered through the relative share of various actions in the current set of strategies. Consequently, individuals are just as likely to mutate to an action previously tested with very uncomfortable results as to mutate to an action that they have never tried before. In reality and in learning models such as fictitious play and Bayesian learning, people would remember their past experience and would treat the two actions differently.

In the field of technical optimisation, genetic algorithms and evolutionary strategies are still used for different applications, contrastingly, computational economists only use genetic algorithms. The literature does not provide a reason for this neglect of evolutionary strategies. Originally there has been one major difference between the two approaches: evolutionary strategies require the variables that are to be optimised to be real values, while genetic algorithms require a binary coding. This had, of course, some consequences for the modelling of mutations and crossovers. In the case of genetic algorithms mutations are switches of bits in the binary code, while in evolutionary strategies mutations are normally distributed changes in these real values (see Rechenberg 1973 and Schwefel 1995). Similarly, crossovers, in the case of genetic algorithms, are the exchange of bits, while crossovers are used in a similar form in evolutionary strategies only if a multi-dimensional variable is to be optimised. However, crossovers in the case of genetic algorithms should also be only used if the bits represent independent features of behaviour. If the bits in a binary string of a genetic algorithm represent the binary coding of a value, crossover might lead to strange results because the crossover of 1000 (representing 8) with 0111 (represent-

ing 7) might lead to 1111 (15) and 0000 (0), which is not in line with the interpretation of crossovers as representing communication.
Hence, the coding of variables is the basic difference between genetic algorithms and evolutionary strategies. Therefore, in technical optimisation, which of the two approaches are used depends on which coding is more adequate for the given problem. In contrast, in economics only genetic algorithms are used and recently authors have started to adapt genetic algorithms to the use of real values instead of the binary coding. It seems as if computational economics are simply throwing away half the available options.

3.7 Combined models: EWA and VID model

There seems to be a natural tendency for researchers to search for general models that describe various processes, which also holds in the learning context. We have seen above that various kinds of learning processes exist and it could be argued that the different models can be combined. Combining different models of routine-based learning can be justified by the argument that each routine-based learning model represents one feature of learning and that all these features are simultaneously given. However, the modelling of routine-based learning was justified by the attempt to focus on one feature to simplify modelling. This justification is lost by combining routine-based learning models. Nevertheless, combined learning models exist and two approaches are presented here: Camerer and Ho's Experience-Weighted Attraction (EWA) model and the Variation-Imitation-Decision (VID) model.
In the EWA model (Camerer & Ho 1999) it is argued that two fundamental types of learning processes exist: reinforcement learning and belief learning. The model is designed such that it describes these two learning processes as border cases for specific choices of the model parameters. The model is described by two equations that determine the process of updating in the light of new experience:

$$N(t) = \rho \cdot N(t-1) + 1 \tag{15}$$

and

$$A_i^j(t) = \frac{\phi \cdot N(t-1) \cdot A_i^j(t-1) + [\delta + (1-\delta) \cdot I(s_i^j, s_i(t))] \cdot \pi_i(s_i^j, s_{-i}(t))}{N(t)} . \tag{16}$$

$N(t)$ is called the experience weight and $A_i^j(t)$ is called the attraction of strategy j for individual i. The term $s_i(t)$ denotes the strategy used by individual i at time t, while $s_{-i}(t)$ is a vector that represents the strategies that are chosen by all other individuals, except individual i, at time t. The function $I(s_i^j, s_i(t))$ equals one if $s_i^j = s_i(t)$ holds and equals zero otherwise. The payoff $\pi_i(s_i^j, s_{-i}(t))$ is obtained by individual i if she chooses strategy s_i^j and the behaviour of all others is described by $s_{-i}(t)$. The terms ρ, ϕ, and δ are the parameters of the model. The initial values of $N(t)$ and $A_i^j(t)$ have to be chosen according to considerations about what experience individuals might transfer from other situations.
If $N(0) = 1$ and $\rho = \delta = 0$, the model reduces to the original Roth-Erev model of reinforcement learning. If δ is larger than zero, the experience collection is expanded to

actions that are not chosen. Thus, it is assumed that the individual can learn by the observation of events about the adequacy of actions that were not taken. If $\rho = \phi$ and $\delta = 1$, the model reduces to weighted fictitious play. For other parameter values the model presents a mixture of the two learning processes. Finally, the EWA model assumes that people make their choice of action according to a logit formulation. The probability of each strategy j to be taken by individual i is given by

$$P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}} \tag{17}$$

where λ is a parameter and m_i is the number of possible strategies that individual i can use.

By combining a reinforcement learning model and a belief learning model, the EWA model presents a mixture of non-conscious and conscious learning. Hence, it could be argued that the EWA model is a general learning model.

The VID model combines all the features of learning that have been described above under the heading of routine-based learning: experimentation, experience collection, imitation, and satisficing. Therefore, it presents what has been discussed above: a combination of all main features of conscious learning on the routine level. However, this implies that the VID model is very complex, which reduces its attractiveness. In most situations, the model would contain many aspects that are simply irrelevant. Hence, its adequacy is restricted to addressing some general questions (see, e.g., Brenner 2001).

A complete description of the model would take too much space here. Therefore, the interested reader is referred to the detailed description by Brenner (1999, Ch. 3). The model assumes that individuals collect information on the outcome of behaviours similar to how it is done in fictitious play. However, this knowledge does not directly influence behaviour. Instead, individuals are assumed to continue displaying the same behaviour if it leads to satisfying outcomes most of the time. Some very rare modifications are assumed without motivation. Besides these, individuals change behaviour only if they are dissatisfied with the previous outcomes of their actions. In this case they choose their next action according to their experience and their observation of others.

The resulting model contains many parameters and is a mixture of various existing models and some psychological findings. It is unclear whether the way in which the model is put together is a realistic one. Many other ways of building such a combined model of routine-based learning are possible. However, we have stated above that routine-based models are not developed in order to describe reality exactly. They are designed to represent certain features of learning processes approximately. In this way, the VID model could be used to show the consequence of combining various learning features.

4. Modelling belief learning

The psychological literature on learning processes is nowadays dominated by cognitive learning process analysis. Neuro-scientific research has added quite some insights to this

line of research. However, the class of belief learning does not exactly match what psychologists call cognitive learning. It is rather a subclass of cognitive learning processes.
Nevertheless, the discussion of conscious learning will be based on psychological knowledge. Therefore, we start with an overview on the psychological knowledge on cognitive learning. Then, the models used in economics are explicitly described. We subsume under this category not only the learning models that are called belief learning models in economics, but also the rational learning models as well as many models from artificial intelligence and machine learning. However, despite this large number of available models there is little empirical evidence from experimental economics.

4.1 Psychological findings about cognitive learning

At the beginning of the psychological research into cognitive learning, the main issue was the development of so-called cognitive or mental models or maps (see, e.g., Bruner 1973, Piaget 1976, Johnson-Laird 1983 and Anderson 2000). They are based on the argument that within the brain a representation of the real world (or at least the part of the real world that is relevant to the individual) is developed. This representation contains subjective knowledge about concepts, connections, causal relationships and so on. The representations in the brain develop according to experience and information obtained from different sources. Still, psychology has not come to one common framework to deal with these processes and many questions are left unanswered. Therefore, we present here one basic concept in order to build some ground for the following discussion. Then, we shall present those insights about cognitive learning that are most relevant to our topic of modelling learning as realistically as possible. These insights will be drawn from different sources.
The concept explained here is that of mental models. This term is chosen for examination as it is the only one of the available concepts which has been introduced to economics (Denzau & North 1994) and has frequently been used thereafter (see, e.g., Gößling 1996 and Kubon-Gilke 1997).
In psychology the theory of mental models has mainly been influenced by Johnson-Laird (1983). The basic idea is that individuals develop mental models about their surroundings.
A detailed model on the development of causal relations according to this theory is provided by Goldvarg & Johnson-Laird (2001). Denzau and North, who have introduced this concept into economics, state that "mental models are the internal representation that individual cognitive systems create to interpret the environment" (Denzau & North 1994, p. 4).
Hence, we can interpret mental models as the sum of all beliefs and knowledge that an individual holds about the world, including the results that different actions will bring about. Mental models are subjective and may not match reality.
Mental models about the working and state of the real world are used to make predictions about the future and the consequences of actions. This, in turn, is the basis for choosing an adequate action. A description of this process should be based on those mechanisms that guide the development of mental models. However, these are difficult to study. From the research into neural networks, we know that these networks are able to reproduce very complex relationships. The structure of the neural network in human brains is more complex than the neural networks that are usually implemented on computers. Hence,

people are able to develop extremely complex mental models.
Furthermore, new information is always interpreted and included in the light of existing mental models. The subjective knowledge of an individual is somehow structured in hierarchies (psychologists are still discussing their exact structural appearance). Each new piece of information may change different hierarchical levels. This could cause complex and elaborated mental models. However, experimental studies show that people usually consider only a few levels of a strategic situation (see, e.g., Nagel 1995 and Stahl 1998). Hence, most parts of mental models can be assumed to have a rather simple structure.
Mental models contain various elements that are labelled differently by various researchers. In some cases it is helpful to restrict the discussion to specific parts of mental models. Here the categorisation of Anderson (2000) is used. He distinguishes between propositions, schemas and scripts. Propositions represent "what is important about specific things" (Anderson 2000, p.155). Schemas sort similar things together and define what they have in common. Scripts are representations of events and sequences of events. Thus, scripts also represent what the individual expects to happen under certain circumstances. Decisions are made on the basis of these expectations and become habits if the same situation is repeatedly faced and the decisions prove to be adequate. Scripts, and thus also expectations, change if a new experience is made. However, people are reluctant to change scripts. They are much faster in processing confirming evidence for the existing scripts than in processing experience that does not fit into the existing scripts and schemas (see, e.g., Kahneman 2003 and Hebb 2002 for the neurological basis of this process). Scripts are the elements that are of most interest in the context of belief learning. Some findings that are relevant for modelling belief learning are:

- People typically hold one mental model about reality at any one time (see Dörner 1999). Sometimes an individual might not be sure about certain issues and may consider different expectations. However, people tend to fix their expectations quickly on the basis of little evidence.

- Scripts, and hence also expectations, change if new knowledge about a situation is gathered (see Anderson 2000). New knowledge can be obtained by experience, observation or communication.

- Experiments have shown that people do not develop very complex expectations (see Stahl & Wilson 1994). However, if a situation is repeatedly faced and simple expectations are falsified, people develop more complex expectations (see Brenner & Hennig-Schmidt 2005).

- People develop scripts quickly without much evidence and tend to stick to scripts without strong opposing evidence (see Dörner 1999). People have the ability to ignore evidence that contradicts their beliefs.

4.2 Fictitious play

The fictitious play model was developed within the context of games (see Brown 1951). It assumes that individuals in a game mentally record all previous moves of their opponents.

Let us denote each move of their opponents by the vector $\mathbf{a}_{i\_}(t)$ and their own action by $a_i(t)$. The individuals are assumed to remember all previous behaviours of all other individuals. Thus, they are able to calculate the frequency of occurrence for each action profile $\mathbf{a}_{i\_}$. They assume that their opponents' actions will occur with the same probability in the future. Consequently, the expected probability $p(\mathbf{a}_{i\_}, t)$ for each action profile $\mathbf{a}_{i\_}$ realised by the other individuals is given by

$$E\Big(p(\mathbf{a}_{i\_}, t)\Big) = \frac{1}{t} \sum_{\tau=0}^{t-1} \delta(\mathbf{a}_{i\_}(\tau) = \mathbf{a}_{i\_}) \tag{18}$$

where

$$\delta(\mathbf{a}_{i\_}(\tau) = \mathbf{a}_{i\_}) = \begin{cases} 1 & \text{for} \quad \mathbf{a}_{i\_}(\tau) = \mathbf{a}_{i\_} \\ 0 & \text{for} \quad \mathbf{a}_{i\_}(\tau) \neq \mathbf{a}_{i\_} \end{cases} . \tag{19}$$

Furthermore, the individuals have complete knowledge about their payoffs $\Pi_i(a_i, \mathbf{a}_{i\_})$ for each action profile $(a_i, \mathbf{a}_{i\_})$. So they are able to calculate the best response to the expected behaviours of their opponents. To this end, they calculate the expected average payoff

$$E(\Pi_i(a_i, t)) = \sum_{\mathbf{a}_{i\_}} \Pi_i(a_i, \mathbf{a}_{i\_}) \cdot E(p(\mathbf{a}_{i\_}, t)) \tag{20}$$

for each action a_i they are able to realise. Then they choose the action a_i with the highest expected average payoff $E(\Pi_i(a_i, t))$. This action is called the best response to the expectations given by $E(p(\mathbf{a}_{i\_}, t))$.

Of course, the above model is not restricted to learning the behaviour of other players in a game. It can also be applied to any situation in which individuals learn about the frequency of certain events, are able to observe these events after their own action, and know the impact which these events, in combination with their own action, have on their payoff or utility. Hence, the fictitious play model can be easily applied to the learning of beliefs. All that a researcher has to do is to define the set of events and/or causal relations that the individuals build beliefs about.

As modelled above, fictitious play assumes that the likelihood of these events and causal relations is given by a stationary probability distribution. Hence, all the individuals have to do is to approximate this probability distribution by collecting more and more information and calculating appropriate averages. This is done in Equation (18). If the real probabilities change, the fictitious play model allows only a very slow adaptation to the new circumstances. The more individuals have already learned, the less flexible their expectations become. If circumstances continually change, the above fictitious play model is, of course, a rather incompetent learning method.

Moreover the fictitious play model, as it is described above, requires an enormous cognitive capacity because all previous experience has to be remembered and consequently the best response calculated. It is doubtful whether individuals are able to do so.

In recent years, some modifications of the fictitious play model have been presented. These modifications reduce the requirements for the individuals' cognitive capacity. Young (1993) modelled individuals who are only able to remember the last k events. They play the best response based on the average observations in these k rounds. By doing so, the individuals

adapt faster to changing circumstances. If k is reduced to one in Young's model, the model of myopic learning is obtained (see, e.g., Ellison 1993, Kandori, Mailath & Rob 1993 and Samuelson 1994). However, this extreme again seems to be less realistic.
Another more realistic possibility is to exponentially weigh past experiences (see also the description of the model by Cheung and Friedmand in Duffy 2005). This means that the latest experience is more important than the experience that occurred further back in the past. For such a model Equation (18) has to be replaced by

$$E\Big(p(\mathbf{a}_{i_-},t)\Big) = \frac{1-\beta}{1-\beta^{(t-1)}} \sum_{\tau=0}^{t-1} \beta^{(t-1-\tau)} \cdot \delta(\mathbf{a}_{i_-}(\tau) = \mathbf{a}_{i_-}) \tag{21}$$

where β is a parameter that determines how fast experience is forgotten.
Some authors have modified the concept of myopic learning by the introduction of errors and occasional adaptation to the best response (see, e.g., Samuelson 1994) or by the introduction of gradual convergence to the best response behaviour (see, e.g., Crawford 1995). All this makes the model less demanding with respect to peoples' cognitive capabilities and thus more realistic. Nevertheless, all these versions of fictitious play share the assumption that individuals define a set of events or relationships and keep track of their likelihood.

4.3 Bayesian learning

Although very few economists would claim that people behave optimally, most of economic literature on learning has a connection to optimisation. The oldest and most prominent 'optimal' learning model is Bayesian learning (descriptions and analyses of Bayesian learning can be found in Jordan 1991, Eichberger, Haller & Milne 1993, Kalai & Lehrer 1993, Jordan 1995 and Bergemann & Välimäki 1996). Bayesian learning concerns a single learning individual and assumes that the individual establishes a set of hypotheses about the situation she faces. Each hypothesis h makes a probabilistic statement $P(e|h)$ about the occurrence of each event e of a set of events $\mathcal{E}$. This means that hypothesis h implies that event e occurs with probability $P(e|h)$. The set of hypotheses $\mathcal{H}$ has to be complete and complementary, meaning that every possible state of reality has to be represented by one, and only one, hypothesis. At the beginning of the learning process an individual generally assigns the same probability $p(h,0)$ to each hypothesis $h \in \mathcal{H}$. If the individual has initial information about the situation she faces, the initial probabilities $p(h,0)$ are different from each other according to this information. $p(h,t)$ denotes the individual estimation of the probability that hypothesis h is correct. In other words, $p(h,t)$ is the belief of the individual in hypothesis h at time t. The sum $\sum_{h\in\mathcal{H}} p(h,t)$ has to equal one.
After each event $e(t)$ the individual updates her presumed probabilities. The updating proceeds as follows (cf. e. g. Easley & Kiefer 1988 or Jordan 1991). The individual calculates the probability $P(e(t)|h)$ for each hypothesis h. Subsequently she updates her beliefs according to the following equation:

$$p(h,t+1) = \frac{P(e(t)|h) \cdot p(h,t)}{\sum_{\tilde{h}\in\mathcal{H}} P(e(t)|\tilde{h})\; p(\tilde{h},t)}\ . \tag{22}$$

By this, the presumed probabilities of hypotheses that predict the occurrence of the observed event with a greater chance increase, while the presumed probabilities of the other hypotheses decrease. The condition $\sum_{h \in \mathcal{H}} p(h,t) = 1$ is maintained while updating the probabilities according to equation (22). After many observed events, the probability $p(h,t)$ should converge to $p(h,t) \approx 1$ for the correct hypothesis about reality, and to $p(h,t) \approx 0$ for all other hypotheses.

Decisions are made according to the following consideration. For each hypothesis h the individual calculates the average utility $\bar{u}(a,t)$ that the action a gives rise to. To this end, she has to assign a utility $u(e,a)$ to each event e and action. $\bar{u}(a)$ is given by

$$\bar{u}(a,t) = \sum_{h \in \mathcal{H}} \sum_{e \in \mathcal{E}} u(e,a) \cdot p(h,t) \cdot P(e|h) \; . \tag{23}$$

The average utility is the expected result from action a. In economics it is called expected utility. As it is learnt adaptively, it is also referred to as adaptive expected utility. Subsequently, the individual decides in such a way that she maximises her expected utility $\bar{u}(a,t)$.

4.4 Least-squares learning

Another learning model that is based on the assumption that people optimise their behaviour is least-squares learning (see Bray 1982, Marcet & Sargent 1989, and Bullard & Duffy 1994). In this model, it is assumed that people make assumptions about the functional dependencies in reality. These dependencies contain, as in regression analysis, a number of parameters. Individuals, it is assumed, intend to learn about the value of these parameters. In order to predict the values of these parameters it is assumed that they proceed statistically. It is further assumed that individuals fit the parameters such that the sum of the squares of the differences between the predicted and the observed values becomes minimal.

If an individual, for example, assumes a linear relationship between $y(t)$ and $\tilde{y}(t)$, the slope parameter β in this linear function can be calculated by

$$\hat{\beta}(t+1) = \frac{\sum_{t'=1}^{t} y(t')\,\tilde{y}(t')}{\sum_{t'=1}^{t-1} y^2(t')} \; . \tag{24}$$

$\hat{\beta}(t+1)$ is the prediction of β at time $(t+1)$. The formula for linear regression can also be written recursively as follows:

$$\hat{\beta}(t+1) = \hat{\beta}(t) + g(t)\Big(\frac{\tilde{y}(t)}{y(t)} - \hat{\beta}(t)\Big) \tag{25}$$

and

$$g(t) = \left(\frac{y^2(t)}{y^2(t-1)\; g(t-1)} + 1\right)^{-1} , \tag{26}$$

where $g(t)$ exists only for mathematical reasons and has no economic meaning.

The decision is then made on the basis of the estimated value. In the long run, the algorithm converges to the real value of β if this value is constant (cf. Marcet & Sargent 1989).

4.5 Genetic programming

Genetic programming has emerged from the concept of genetic algorithms (see Bäck 1996 for a description of all types of evolutionary algorithms). The basic mechanisms of genetic programming are the same as those of genetic algorithms: selection, reproduction, crossover and mutation. The difference is the unit that is selected and mutated. In the case of genetic algorithms actions or strategies are coded, usually in binary form, and optimised by the algorithm. In the case of genetic programming a formula- or program-like structure is coded and optimised. This formula- or program-like structure can be easily interpreted as a belief about the functioning of the world.

A usual example is the coding of a mathematical formula. For example, the formula $y = 3 \cdot x_1 + 8 \cdot (x_2 - 1)$ would be coded in genetic programming as depicted in Figure 1. It might be assumed that it represents the belief of an individual about the relationship between the variables x_1, x_2 and y. Such a representation allows economic agents to have quite complex beliefs. Furthermore, the beliefs are not restricted at the beginning by the structure of the formula as in the case of fictitious play. Therefore, genetic programming seems to be adequate to describe belief learning (for a similar argument see Chen, Duffy & Yeh 2002). If, furthermore, the formula is length restricted, the psychological finding that people tend to think in simple relationships is included in the modelling.

The learning process is modelled in genetic programming by the processes of selection, reproduction, crossover and mutation. At each point in time a number of formulas or programs coded as given in Figure 1 exist. Some of these are selected according to the correspondence between their prediction and the observations of the real world (the sum of squares of the errors can be used as in least-squares learning). These selected formulas or programs are reproduced. Then cross-over operations are applied as in genetic algorithms. To this end two formulas or programs are crossed together at a randomly determined node and the two parts connected by this node are exchanged. Finally, the resulting formulas or programs are mutated (for a detailed description see Bäck 1996).

4.6 Classifier systems

The psychological literature often characterises human beings as classifiers. Humans tend to sort things, events and relationships into classes and act according to their classification. Hence, it seems to be natural to model such classifying behaviour.

Classifier systems seem to be an adequate tool for this purpose (see Holland, Holyoak, Nisbett & Thagard 1986). The core elements of classifier systems are condition-action rules. These rules state under what conditions specific actions should be taken. Thus, two things have to be codified. First, the set of conditions has to be defined, which is usually, but not necessarily, in binary form. However, it is generally in the form of a string of characteristics: $\{c_1, c_2, ..., c_n\}$. The same holds for the actions: $\{a_1, a_2, ..., a_p\}$.

A classifier system is characterised at each time by a set of q decision rules R_i $(i = 1, 2, ..., q)$ of the form $\{c_{i1}, c_{i2}, ..., c_{in}\} \rightarrow \{a_{i1}, a_{i2}, ..., a_{ip}\}$. In this context each entry in the condition string can be represented by a symbol '#' instead of a number, which implies that the corresponding action is taken independent of the value of this characteristic.

Two values are assigned to each decision rule R_i at each time: its strength, which is determined by its success in the past, and its specificity, which is determined by the number of '#'s in the condition string. If a message $S = \{s_1, s_2, ..., s_n\}$ that characterises the current situation is observed, this message is compared to all condition strings. Decision rules with a condition string that matches S compete for being activated. The value

$$B(R_i) = g_1 \cdot (g_2 + g_3 \cdot \text{Specificity}(R_i)) \cdot \text{Strength}(R_i, t)$$

(g_1, g_2 and g_3 are fixed parameters) is calculated for each of these decision rules. The decision rule with the highest value is activitated and the respective action taken. The strength of this decision rule is updated according to

$$\text{Strength}(R_i, t + 1) = \text{Strength}(R_i, t) + \text{Payoff}(t) - B(R_i) \ .$$

The specification of classifier systems has so far created mechanisms to identify the most adequate rules out of a given set of rules. However, no new rules evolve. Therefore, the existing decision rules are changed by a second process, which is based on genetic operators. At certain points in time, a certain number of decision rules are eliminated. The probability of each decision rule being eliminated decreases with the actual strength assigned to the rule. To replace the eliminated rules some new rules are created. To this end some existing rules are randomly picked. The probability of each rule to be picked increases with its actual strength. These rules are copied and then slightly modified. Different specifications of these mechanisms exist in the literature (see, e.g., Beltrametti, Fiorentini, Marengo & Tamborini 1997).

4.7 Neural networks

In the last decade, computer technology has developed to such a point that reproducing brain structures on the computer has to some extent become feasible. Hence, it seems to be natural to model human cognitive learning processes by simply rebuilding the brain on the computer. This is done in the field of neural networks.

Many discussions of neural networks can be found in the literature so that it is unnecessary to discuss them at length here (see, e.g., Baltratti, Margarita & Terna 1996). They have been used repeatedly in the recent economic literature to model learning processes (see, e.g., Calderini & Metcalfe 1998 and Heinemann 2000). Nevertheless, the prominence of this method of modelling learning processes is still slight. There are two main reasons for this. First, the details of how brain structures are developed and how meaning is created within these networks are not sufficiently known. As a consequence, it is difficult to determine how a neural network that rebuilds the human brain has to be designed. Second, using a neural network, which needs to be quite complex, does not allow us to understand why the modelled agent behaves in a certain way. Using neural networks is akin to a black-box approach. The results of such an approach are difficult to judge as one cannot be sure that the network has been adequately designed.

4.8 Rule learning

Some psychologists have claimed that cognitive learning follows also the rules of reinforcement learning (see, e.g., Kandel & Schwartz 1982). The only difference is that rules are

reinforced instead of actions. This is taken up in the approaches by Erev, Bereby-Meyer and Roth (1999) and Stahl (2000), the latter calling this process rule learning.
This means that a probability is assigned to each alternative belief or script. Each new experience changes the probabilities according to the mathematical formulation of reinforcement learning. The usual modelling of reinforcement learning implies that if a decision has to be made, a belief is randomly drawn according to the probabilities and the action is taken that is most suitable given this belief. Such a modelling might be adequate if individuals are not consciously aware of their expectations and behave intuitively, but are nevertheless tacitly guided by beliefs and scripts. If individuals are aware of their beliefs it seems unlikely that decisions are made according to one randomly drawn belief each time.

4.9 Stochastic belief learning

A similar approach that takes psychological findings into account is the stochastic belief learning model (see Brenner 2005). This model differs from rule learning, and also Bayesian and least-squares learning, especially in two features. First, it assumes that not all possible beliefs play a role. Instead, a set of relevant beliefs is defined according to experimental knowledge. Second, it assumes that individuals only consider one belief most of the time.
The set of possible beliefs is denoted by $\mathcal{H}$ and each element is denoted by $h \in \mathcal{H}$. The number of different beliefs is denoted by H. The beliefs of each individual i at any time t are given by a set of beliefs $s_i(t) \subset \mathcal{H}$ which is a subset of all possible beliefs. This means that each individual considers at any one time only a few beliefs. The model that is proposed here starts from a situation in which each individual holds exactly one belief, meaning that $s_i(0)$ contains exactly one element denoted by $h_i(0)$ here. The initial beliefs of individuals have to be empirically determined. Usually, beliefs from other situations are transferred. However, so far, there is no available knowledge on this process that would allow us to make predictions about the initial beliefs.
The beliefs are then updated according to incoming information. This information might originate from one's own experience, from observing or communicating with others. The information is only used to update beliefs that are currently considered by an individual. Each belief h in the set $s_i(t)$ is checked against the new information k ($\mathcal{K}$ denotes the set of all possible pieces of information that might be gained). Only two situations are distinguished here: the new information might either contradict or not contradict the belief h.
If none of the beliefs in $s_i(t)$ is contradicted by the new information obtained at time t, the set of beliefs remains unchanged: $s_i(t+1) = s_i(t)$. For each belief h in the set $s_i(t)$ contradicted by the new information k, it is randomly determined whether this belief disappears from the set $s_i(t+1)$. According to psychological arguments above, a belief is not automatically eliminated if it is contradicted or proven wrong. People tend to stick to their beliefs even if there is conflicting evidence. Hence, a probability ρ_i is defined for each individual i that determines the likelihood of a belief h to be eliminated in face of information k that contradicts this belief. Hence, ρ_i describes how individual i reacts to new knowledge. The smaller ρ_i, the more individual i sticks to her beliefs.
According to the above process, beliefs could disappear. In contrast, new beliefs appear

according to three processes: variation, communication and necessity. First, the model assumes, with a certain likelihood, that individuals consider a new belief by chance. The probability that a new belief is considered by chance is denoted by ν_i. Then, this new belief is added to the set $s_i(t+1)$. Second, an individual might be convinced to consider a belief by others. This can be modelled by assuming that each individual i communicates the beliefs in her set $s_i(t)$ at time t to each other individual j with a certain probability σ_{ij}. σ_{ij} would then describe the probability that at each time t the elements of $s_i(t)$ move into the set $s_j(t+1)$. Third, if a set $s_i(t+1)$ is empty at the end of time t, a new belief has to be taken up. There always has to be at least one element in the set $s_i(t)$ of beliefs, because otherwise the individual is unable to decide about her action.
If, for one of the above three reasons, a new belief is built, this new belief is determined as follows. Each belief h that individual i does not hold so far ($h \in \mathcal{H} \backslash s_i(t)$) is chosen with a probability that is given by

$$P_i(h,t) = \frac{p_i(h,t)}{\sum_{\tilde{h} \in \mathcal{H} \backslash s_i(t)} p_i(\tilde{h},t)} \; . \qquad (27)$$

This means that the values $p_i(h,t)$ determine the likelihood of each belief h to be considered. The initial values of these probabilities have to be empirically estimated or it has to be assumed that each belief is equally likely.
During the learning process it can be assumed that beliefs that have been considered in the past and are then omitted because of contradicting events are less likely to be reconsidered. Hence, the model assumes that each time a belief h leaves the set $s_i(t)$, the probability for this belief to be reconsidered is updated according to

$$p_i(h,t+1) = \lambda_i \, p_i(h,t) \; . \qquad (28)$$

where λ_i is a parameter that determines how likely individual i reconsiders disconcerned beliefs.

5. Conclusions and Recommendations

Each kind of learning process; non-conscious learning, routine-based learning and belief learning; has been discussed separately above. Now, we will return to the general question of choosing a learning model for conducting a computational study.
The use of learning models is often criticised due to the lack of a common model and the ad-hoc choice of specific models. It has been argued above that a common model is unlikely to exist as different learning processes take place in different situations. However, learning models are indeed usually chosen without much justification. This section aims to offer a common platform to justify the use of specific learning models in specific contexts.
The recommendation consists of a two-step process. First, it has to be decided which type of learning should be modelled in the given context. The characteristics of the situation determine whether economic agents learn non-consciously or consciously and whether routine-based modelling is sufficient (Section 5.1).
In a second step, a learning model has to be chosen within the relevant class of learning. There are different ways to choose a learning model and various information sources

that can be used (Section 5.2). Here, one specific approach to choose a learning model is taken and the above presented learning models are discussed on the basis of this approach (Section 5.3). However, some degree of freedom remains as the lack of empirical and experimental evidence makes it impossible to precisely recommend one learning model for each learning class. Researchers have to make their final choice according to the specific topic of their study and to some extent according to their own preferences. The whole process of choosing a learning model is summarised in Figure 2.

5.1 Situational characteristics and learning

A very important topic that is rarely discussed in the literature concerns the applicability of learning models in various circumstances. Above, three types of learning have been distinguished: non-conscious learning, routine-based learning and belief learning. Furthermore, it has been argued that basically there are two different learning processes; non-conscious learning and belief learning; while the third type, routine-based learning, presents a simplification of belief learning. Hence, the first question to be answered is when do non-conscious learning and belief learning occur in reality; and the second question addresses when belief learning can be appropriately approximated by routine-based learning models.

5.1.1 Non-conscious versus belief learning

As mentioned above, non-conscious learning seems to be a hard-coded process that takes place in many, if not all, animals. Examples in the economic sphere are affective purchases, tacit knowledge, intuition and routines of interaction within and among firms. In contrast, we will usually not buy a car or a house in an affective way. Humans are able to reflect on their behaviour and build models about the consequences. Such conscious learning seems to be capable of reducing the consequences of non-conscious learning.

As a consequence, non-conscious learning is, in general, mainly relevant if conscious learning does not take place. Conscious learning requires the individual to be aware and reflect upon their behaviour. Therefore, it requires time and cognitive capacity. During a normal day we face an enormous number of decision-making tasks of varying degrees of importance and difficulty. Most decisions are made automatically, without spending a single thought on them. An obvious example is driving a car on the left- or right-hand side of the road. In a familiar location this is luckily an unconscious decision. However, many people have experienced this non-conscious behaviour in a foreign country, where people drive on the other side of the road.

Humans seem to be endowed with a mechanism that determines which decisions and behaviours are consciously reflected. This mechanism is not studied in the literature. Hence, we can only speculate about its functioning. Some statements seem to be evident. Conscious reflection on behaviour is restricted by the time available for such contemplation. Furthermore, there are certain motivations that make conscious reflection more likely. Finally, habituation can also occur. All in all, this means that once a situation has been consciously reflected on and a conclusion has been drawn, individuals often repeat the same action without further reflection.

In general we can state that people spend their cognitive time on those behaviours and decisions that they personally find most important to consider. All other actions are the result of non-conscious learning. Various situations provide obvious motives provoking conscious attention:

- In a new situation in which individuals have no established rules to rely on, a cognitive effort is worthwhile as an arbitrary choice may cause poor performance. Nevertheless, in many new situations individuals utilise routines transferred from similar situations. Thus, the expression 'new' has to be handled carefully.

- Dissatisfaction is a strong motivation for dealing with a situation cognitively. When a repeatedly faced situation leads to unsatisfactory outcomes, individuals are motivated to change their behaviour. To this end, they are attentive to the situation and reflect upon it cognitively, in an attempt to improve their performance. In this context the aspiration level plays an important role.

- Certain situations and decisions are regarded as important to the individual for personal reasons, such as personal pride, aesthetic aspirations or the relevance for one's own life. Such specific motivations may explain the large variance of behaviour among individuals.

Except the latter point the above arguments state that attention is directed towards situations in which the outcome is either unsatisfactory or can be anticipated to be unsatisfactory with a high probability. Thus, we claim the following: in general, non-conscious learning is the normal route to behavioural changes. Individuals generally do not pay attention to the repeated situations they face. Paying conscious attention is inspired by unsatisfactory results or their anticipation. Although this assumption is a theoretical abstraction of the factors analysed above, it offers a sound basis for the categorisation of learning processes and for the choice of a learning model to describe the learning process. Whenever a modeller treats a situation that is important or new to the individuals, it is likely that the individual is learning consciously and a conscious learning model should be chosen. If the situation that is modelled is unimportant for the economic agents, it seems adequate to assume that they learn non-consciously and thus, a non-conscious learning model should be chosen.

In reference to the above simplification, however, it must be noted that in the course of a repeated situation faced by an individual, two non-conscious processes should be distinguished. A behaviour may originate and be guided further non-consciously until an unsatisfactory outcome is obtained, or may originate consciously and then be guided further non-consciously. In the case of consciously learnt behaviour, individuals may, as soon as they have found a satisfying behaviour, direct their cognitive attention to other situations. Subsequently, the behaviour becomes subject to non-conscious learning and, as it is repeatedly chosen, this behaviour is confirmed as long as it is reinforcing, i.e., satisfying. If, however, the outcomes prove dissatisfying, conscious attention is usually redirected to the situation.

5.1.2 Routine-based versus belief learning

The choice between the models of the two kinds of conscious learning, routine-based and belief learning, is more difficult. As mentioned above, routine-based learning does not describe a real learning process but is an approximation of belief learning. Hence, the question is not in which situations the two kinds of learning processes occur but when the routine-based approximation sufficiently describes the real learning process.
Models of routine-based learning are usually based on empirical or experimental observations of learning. They reduce the learning process to one or a few main characteristics that can be observed without actually knowing the real processes responsible for the learning process in the brain. Through this, they miss part of the conscious learning process. The missing part could be called the understanding of the situation. Let us consider as an example the process of imitation. If we imitate other people we do not only imitate the behaviour that performs best, we usually also develop a subjective understanding of why they perform best. This includes the possibility to act differently if the situation changes or if the observed agents differ from the own personality.
Nevertheless, in some situations the development of a deeper understanding does not influence behaviour significantly. If, for example, all people sit in front of identical multi-arm bandits and all want to maximise their profits, not much understanding of the situation is needed and people will imitate the choice of the arm that performs best, possibly with some experimentation of other arms. It would be unnecessarily complex to accurately model the processes in the brain in such a situation.
Hence, the question of whether to use a routine-based or belief learning model is related to the discussion between using a complex and realistic model or whether to use a simple, approximating model. In addition, this question is related to the discussion on the validity of models. These questions are taken up in the next section. In the sphere of routine-based models there is quite an amount of supporting evidence available while in the sphere of belief learning, no model has yet been developed that seems completely convincing. Hence, there is a temptation to use routine-based learning models and therefore these are discussed here separately. Nevertheless, whoever uses them should be aware of the fact that they only represent approximations of real learning processes. Furthermore, in the literature various routine-based learning models can be found and it is important to choose the most appropriate one.

5.2 Choosing a learning model

After a modeller has clarified the type of learning process to be modelled, one must choose the model that best describes this kind of learning. Some discussion is needed on how such a choice should be done. This discussion has to deal with several questions, such as the aim of choosing a learning model, the sources of empirical evidence, the complexity of learning models, and the level on which learning is modelled.

5.2.1 Aims in choosing a learning model

There are various approaches in choosing a learning model four of which shall be discussed herein. First, one might search for the model that best describes real learning processes. This can be done on the basis of experimental findings or psychological knowledge. However, sufficient empirical and experimental knowledge is not always available. Nevertheless, we will follow this approach here.

Second, one might look for some learning model that leads to an outcome which corresponds to known stylised facts without worrying about the details of the learning model. Such approaches are often taken in agent-based computational economics and aim to keep the learning model as simple as possible and as complex or realistic as needed to obtain the correct outcome. Such a modelling is helpful to understand the minimum requirements of learning in a given situation. It is also helpful to classify situations with respect to the competences that are required of the economic agents in these situation. However, such an approach does not give information about how people learn. Studying whether certain learning models predict economic dynamics that are in line with our empirical knowledge only allows us to reject some learning models but it does not confirm the others. There might be other learning models that lead to the same predictions. Sometimes this is omitted in the literature.

Third, some researchers search for learning models that converge to equilibrium, since often equilibrium is predicted by the neo-classical theory or other equilibrium concepts. It is not clear what we gain from such approaches. Our economic surrounding permanently fluctuates and learning is rather important because it allows us to react to these changes and not because it converges to an equilibrium. In specific cases, however, an equilibrium might describe the real world adequately and searching for learning models that converge to this equilibrium is what has been described above as the second possible aim.

Fourth, some researchers aim at developing clever or even optimal learning models. One might even compare the performance of learning models in a given situation in order to make statements about how people should learn. Besides the positive aim of agent-based computational economics, there is a normative aim of testing alternative economic structures (see Tesfatsion 2001) which may be expanded to alternative behaviours. This is a valid aim which is not further considered here. Most of the recent literature on artificial intelligence and machine learning seems to belong to this approach. In general, a tendency has been observed in recent years to borrow methods from other disciplines. These models have become increasingly complex, mixing such features as evolutionary algorithms, classifier systems, fuzzy logic and neural networks. Besides obtaining very competent learning models, some authors seem to believe that the obtained learning models describe real learning without looking at any evidence for this. As discussed above, this does not hold in all cases.

5.2.2 Validity and complexity of learning models

After clarifying that we want to realistically model learning processes, we have to judge how well confirmed learning models have to be before being used in simulations. Clearly it would be preferred that the learning models available are supported by strong empirical evidence. Unfortunately, only a few studies can be found that offer such evidence. We have to live with and use the little evidence that is currently available and hope for more evidence in the future.

There are two sources of this evidence. On the one hand, experimental studies provide us with some information on the suitability of different learning models (see Duffy 2005). On

the other hand, psychological literature also provides us with information about the mechanisms and circumstances involved in learning. We take the position that learning models that are contradicted by experimental findings (in the situation under consideration) or by psychological knowledge should not be used. Given the above aim, this opinion might be confronted with one counter-argument.
Some researchers argue that their study does not mainly aim to identify the implications of a specific learning process. Instead, their main aim is to analyse a certain complex situation and learning is only included to represent the basic dynamics of behaviour. It might be argued that the choice of learning model is not significant under such circumstances. However, this only holds if different learning models predict a similar behaviour. While this is true for some learning models and situations, it is not universally the case. In various situations different learning models predict contrasting behaviours, so that the choice of the learning model might matter tremendously for the study's result.
Nevertheless, some learning models can lead, under numerous circumstances, to quite similar predictions. Furthermore, empirical and experimental evidence on their suitability is often rare, making it difficult to choose. Thus, the first step would be to exclude all models that can be rejected on the basis of psychological knowledge or experimental evidence, then complexity could be used to select among the remaining models. For example, Rapoport, Seale and Winter (2000) states "that the simplest model should be tried first, and that models postulating higher levels of cognitive sophistication should only be employed as the first one fails".
Hence, in total we have a three selection criterion: experimental evidence, psychological knowledge, and simplicity. All three criteria come with advantages and disadvantages that are valued differently by researchers.
Experimental evidence: The primary source for evaluating the existing learning models are empirical and experimental studies. Empirical studies on learning are quite rare. The existing experimental studies that evaluate learning models are presented in Chapter 5 of this book. It was argued previously that the use of learning models that are rejected by experimental evidence should be avoided. However, two remarks are necessary here.
First, there is still the discussion as to what extent laboratory situations can be comparable to real life situations. Experimental situations are usually artificial and often deprived of any context. Therefore, some researchers argue that behaviour is different in these situations. However, it may also be argued that learning models that adequately describe real life learning processes should also suitably describe experimental learning processes, because in both situations the same cognitive apparatus (the brain) is used. Nevertheless, there may well be a difference in learning processes. In this chapter it is argued that different learning processes exist, for example, non-conscious and conscious learning processes. As a consequence, it may be the case that the frequencies with which certain kinds of learning occur in reality and in experiments could differ tremendously. For example, non-conscious learning processes occur frequently in reality because we do not have the time to reflect on all our decisions. Most experiments, in contrast, force the participants to think about their decisions, so that conscious learning processes appear to dominate in experiments. Thus, we argue that the same kinds of learning processes occur in experiments and real life but that their relative importance may differ between the two settings.

This leads to the second remark. All experimental studies include only a limited number of situations and we know that learning processes differ between situations. Results from experimental studies can only be transferred to a situation under the condition that the situations are sufficiently similar. The amount of sufficient similarity is difficult to state, given the lack of experimental studies that attempt to classify situations according to the learning model that fits best (some discussion in this direction is given in Duffy 2005). Furthermore, the very artificial circumstances often found in experiments must be taken into account.

Nevertheless, we believe that in the long term, experiments will be the major way by which to evaluate the various existing learning models and to support their further development. At the moment there are not many studies of this kind available, so that experimental evidence only offers some help in choosing learning models. Further studies have to be conducted in the future and, most importantly, a classification of situations and a relationship between learning models and situations has to be developed. A primary classification is discussed in this chapter. Checking, refining and revising this classification with the help of experimental studies would tremendously advance the modelling of learning in economics. In addition, these experimental studies should take into account the fact that people differ with respect to their learning processes even in the same situation.

Adequateness of the details of the model: In the psychological literature there is an abundance of knowledge on the details of learning processes. This knowledge can be used to evaluate learning models. However, it might be argued that in economics we are not interested in modelling learning process details. Instead, what we need is a model that effectively describes the resulting behaviour. A model that contradicts psychological findings on the learning process might, nevertheless, predict behaviour correctly. Most of the time agent-based computational economists are only interested in the implications of learning processes for economic processes.

However, as discussed above, evidence about the validity of different model predictions are still rare. Due to the lack of empirical and experimental evidence, adequate representation of the learning process might be a better alternative criterion for learning model evaluation. Detailed models in line with psychological findings could be more trusted than models that contradict psychological findings. However, the inclusion of such details increases the complexity of the model. Hence, there is a trade-off between sufficiently representing all learning details and simplifying the model.

Simplicity of the model: Many economists tend to use simple behavioural models. The neo-classical model of optimisation is a good example. Similarly within the field of economic learning, those models with clear rules and a few parameters have been used most frequently (examples are reinforcement learning, least-squares learning, Bayesian learning and fictitious play). There are good reasons for simplifying learning models.

First, the more parameters a learning model has, an increasing amount of empirical or experimental evidence is necessary to estimate the correct parameters. If there is no sufficient empirical or experimental data, a model with more parameters offers, in general, more vague predictions. Second, simpler models can be more easily interpreted. Third, it is argued that economists are not interested in learning process details, but in their implications for organisation, working and the dynamics of economies. Hence, simple

models that capture the basic characteristics of learning processes should suffice.
However, one might also argue in favour of more complex learning models. First, psychological studies suggest that learning processes are complex. The lack of a simple learning model that exactly describes experimental behaviour offers additional evidence for this claim. Second, the progress in computer technology makes it easy to deal with complex models in simulations. Even for very complex learning models simulations can be run with a large variety of parameter values, so that the influence of different parameter values can be studied.
Of course, more complex models require some additional effort. Hence, we have to deal with a trade-off between the effort necessary and the accuracy of the model. There are many economic situations, such as markets, in which the specific set-up of the learning model is not important for the results of the analysis (see Duffy 2005). However, there are also situations, like a prisoner's dilemma, in which the predictions of different learning models vary tremendously (see Brenner 2005). Thus, in some situations the loss in suitability by using a simple model may be minor, rendering the use of a more complex model unnecessary. In other situations important details may be lost by using a simple model. We have so far only little knowledge on the various situations which are of the former and the latter type. In situations which have a clear equilibrium point and no contradiction between the individually and socially optimal state, simple models seem to be sufficient. Situations involving strategic thinking and assorted motives seem to be, in general, unsatisfactorily described by simple models.

5.2.3 Individual and population learning

An important aspect of modelling learning processes is the level of modelling. Two options exist: Either the learning process of each individual is modelled explicitly – meaning that at each point in time the situation of each individual is clearly represented in the model – or the implications of the individual learning processes for the behaviour of a population of individuals are modelled – meaning that the model only represents the shares of various situations of individuals in the population at any point in time – (for a discussion of this in the context of genetic algorithms see Vriend 2000).
Psychological literature deals almost exclusively with individual learning processes, whereas in the economics literature learning models on the individual and population level are used. Again arguments can be put forward in favour of both options, the individual and population level of modelling learning.
The main advantages of modelling learning on a population level is that it simplifies the modelling and one does not have to care about the details of the individual learning process. For example, modelling on the population level and assuming an infinitely large population eliminates the stochastic feature of learning from the analysis. As a consequence, the resulting learning process can be easily treated analytically. An analysis of learning processes on the population level is also often used in experiments. Examining behaviour on the population level permits us to ignore inter-individual differences. Usually learning models on the population level are more straight-forward as only the fundamental dynamics of learning have to be considered. This makes them quite attractive in situations where the modeller is only interested in the implications of learning processes for an

economy consisting of many agents.
However, neglecting details and individual differences comes at a risk as learning process details and individual differences may indeed matter. There are situations, such as a market, where the exact characteristics of the learning process of individuals is not important for the resulting dynamics. But, there are also situations in which various learning models lead to assorted predictions, which could lead to wrong predictions. Recognising this risk, individual learning models are favoured especially in simulation approaches where the complexity of the implemented learning processes is irrelevant.
An alternative is the use of sub-populations. This is the division of people into heterogeneous groups, whereby the individual characteristics are homogeneous within the group. This prohibits us to model a situation where each individual is specific but takes into account partial differences between individuals. It presents a compromised way to study the impact of heterogeneous types of people.
In conclusion, the question of whether learning processes should be modelled on an individual or population level cannot be answered here finally. Modelling on the population level simplifies things, whereas modelling on the individual level increases accuracy. The situation that is to be studied determines how accurate an individual learning model is in comparison to a model on the population level. In some situations the gain is small, while in others it is tremendous. Hence, it depends on the situation whether the effort of individual modelling is necessary.

5.2.4 Calibration of learning models

Most learning models contain a number of parameters. Thus, once a researcher has chosen a learning model, the parameters of the model must be adjusted accordingly. This is especially important for simulation approaches, wherein each simulation only one specific choice of parameters can be used. Unfortunately, the empirical and experimental literature on learning processes provides us with little information about the parameters of various learning models. Furthermore, parameters may differ between individuals, which is rarely considered in experimental studies. Comprehensive, or even sufficient parameter information of various learning models is not available.
How one might deal with this problem depends on the research aim. Above, different research aims have been outlined. Here it is assumed that the simulation approach is used to predict real processes or to obtain detailed knowledge about the implications of learning processes. In such a case we argue that a range should be defined for each of the parameters such that the modeller is quite certain that the possible values lie within this range. It is important to note that empirical knowledge can be used to reduce the range. All parameter combinations within these ranges have to be analysed in order to be sure about the model implications. A Monte-Carlo approach is applicable in this case (see Werker & Brenner 2004 for a detailed discussion of this methodology and Brenner & Murmann 2003 for an application). If empirical data on learning processes outcomes are available, it can be used in a Bayesian approach to further reduce the parameter ranges or to assign likelihoods to each of the model specifications (see Zellner 1971 and Werker & Brenner 2004). Such an approach is very labour-intensive and requires quite an amount of computer time. However, it increases the reliability of the results.

In addition, such a methodology would certainly benefit from further detailed experimental studies that not only compare learning models but also identify the parameters of various learning models that best describe behaviour. Hopefully, more experimental studies of this kind will be conducted in the future.

5.3 Recommendations for the choice of a model

In this section advice is given to computational economists who try to realistically model learning processes. It has been argued above that three considerations are of help in this context: experimental evidence for a model, the psychological knowledge about the details of real learning processes, and the complexity of learning models. Furthermore, it has been argued above that three kinds of learning processes have to be treated separately: non-conscious learning, routine-based learning, and belief learning. The relevant learning models for each of these learning models are discussed in the following.

5.3.1 Recommendations for non-conscious learning

Section 2 discusses three models that could be used to model non-conscious learning. It has been argued that non-conscious learning processes typically do not occur in experiments. Hence, experimental findings on human behaviour should not be used to judge the adequateness of various models in non-conscious learning. The experimental confirmation of the Roth-Erev model (see Roth & Erev 1995 and Erev & Roth 1998) has to be interpreted as a confirmation of this model being able to describe the outcomes of cognitive learning, although economic literature interprets this model as representing reinforcement learning. This discrepancy results from the difference between the definition of reinforcement learning as simple model on the level of behaviour – although Roth and Erev include more complex aspects such as experimentation and forgetting –, while non-conscious learning is more rigidly defined here as a learning process people are not aware of.

For non-conscious learning, findings from animal experiments and habit formation knowledge could be adopted. From this literature we know that learning processes slow down under constant circumstances but could be reactivated by environmental changes. Furthermore, we know that people might completely eliminate actions from their repertoire and that rewards (positive outcomes) are treated differently from punishments (negative outcomes) (see Kahnemann & Tversky 1979). Only one of the models presented in Section 2 captures all these features, this being the generalised Bush-Mosteller model. Therefore, it seems to be adequate to use this model. As long as there are no negative outcomes, the parameterized learning automaton and the Roth-Erev model without experimentation also represent all major features. However, they are not able to deal with negative outcomes.

5.3.2 Recommendations for routine-based learning

Choosing a model of routine-based learning is more difficult and ad-hoc because these are approximations that consider only part of the real learning process. The parts of the real learning process that should be included depends on the modelled situation. Two different strategies have to be distinguished: using a general model that includes many or all characteristics of routine-based learning or focusing on one specific feature.

General models: Above various general models are presented: the combined models, the EWA model and the VID model, and evolutionary algorithms which also represent, given their usual interpretation, a combined model. Experimental evidence is available for some of these models. The EWA model is supported by some evidence (see Anderson & Camerer 2000), while evolutionary algorithms have been repeatedly confirmed (see Duffy 2005), while the VID model has not yet been experimentally tested.
In comparing the three models in the light of what the psychological literature knows about learning, the VID model is the most accurate. It is developed on the basis of such knowledge. The EWA model combines two models, a reinforcement learning model with a belief learning model. As a consequence, it fails to fit into the classification of learning processes used here and it is not clear whether such a combination does actually occur, although there is separate support for both of these in the psychological literature. Evolutionary algorithms do not match psychological knowledge about learning processes. The cross-over processes are not in line with communication knowledge as well as the interaction and mutation process, which is independent of past experience.
The VID model is obviously the most complex of the three models. Hence, all three models have their shortcomings, so that none of them can be recommended without hesitation. Nevertheless, the use of evolutionary algorithms is favoured here. This recommendation is based on the fact that it might sometimes be helpful to describe the results of a learning process on the population level without much interest in the details of individual learning dynamics. Evolutionary algorithms are well supported by experimental evidence on the population level (see Duffy 2005). Furthermore, evolutionary algorithms have the interesting feature of being able to deal with very large sets of actions and strategies and even allow the sets of strategies to increase endogenously.
Nevertheless, two points have to be kept in mind while using them. First, Rechenberg and Holland developed their algorithms as a means to determine optimal solutions to technical problems. Hence, they are developed to describe an optimising search process and not a learning process. Interpreting their dynamics as individual learning processes seems to be inaccurate, since they contradict psychological knowledge on individual learning processes. Therefore, their use is only recommended on the population level here. Second, the problem of coding actions and strategies should not be neglected while using evolutionary algorithms. Evolutionary strategies and genetic algorithms are two options that mainly differ in their coding. Recoding non-binary values binary and using genetic algorithms thereafter, as it is sometimes done in the literature, seem to be inadequate since it increases the distortion between psychological knowledge on learning and the dynamics of the resulting model.
Separate modelling: All of the above separate routine-based learning models are subject to two limitations. First, they model only one part of the whole learning process. Second, they model the outcome of the underlying learning process, meaning that they only offer an approximation of what actually occurs. There is little knowledge about how suitable this approximation is. Fortunately, there is some experimental evidence in favour of melioriation learning (see Herrnstein 1970), fictitious play (see Duffy 2005), satisficing (Stahl & Haruvy 2002) and some mixed evidence for imitation (Huck, Normann & Oechssler 2002 and Bosch-Domenech & Vriend 2003). There is also various experimental evidence in

favour of the learning direction theory (Berninghaus & Ehrhart 1998 and Grosskopf 2003). However, this theory has very weak predictions and is, therefore, usually not sufficiently precise for agent-based computational models.
The problem with these models is that they only offer accurate approximations of the real learning processes if the learning part described dominates the learning process. With the help of experiments, knowledge on the dominating parts of learning processes in various situations can be established. However, this has not yet been sufficiently done. It depends on the researcher's good judgement to identify which part of learning dominates the examined situation and to choose the respective model.
Learning by experience is not a central research field in psychology. Since melioration learning and fictitious play are both quite simple learning models and are both supported by experimental evidence, they are recommended here. In the case of imitation psychological literature discusses a cognitive process in which people transfer their observations to their own situation. Therefore an extension of the fictitious play model to include imitation is recommended here, although standard imitation models could also be chosen for simplicity. Satisficing is also less prominently discussed in psychological literature but is supported by experiments. How satisficing should be modelled in detail is less clear.

5.3.3 Recommendations for belief learning

Many different models have been described that are potential candidates for modelling belief learning. If we look for experimental evidence, there is some experimental evidence in favour of fictitious play (see Duffy 2005), genetic programming (see Chen, Duffy & Yeh 2002) and stochastic belief learning (see Brenner 2005). Rule learning is somewhat supported by experimental evidence in favour of reinforcement learning. However, the models have never been empirically compared and the experimental evidence has never overwhelmingly favoured any model. For example, the study of Nyarko and Schotter (2002) shows that people hold much stronger beliefs than predicted by fictitious play and change them more radically. Brenner (2005) shows that the stochastic belief learning model explains some individuals' behaviour very well while it fails to explain the behaviour of others.
Furthermore, it is difficult to observe what people are thinking while making their decisions, or as Cheung and Friedman (1997, p. 49) put it: "experimental learning models must deal with the fact that beliefs are not directly observable". This may change in the future as more and more methods are developed to observe people's beliefs in experiments (see Brenner & Hennig-Schmidt 2005 for a promising method). So far, we have to deal with the problem that little evidence is available in the context of belief learning.
Neural networks are in line with the knowledge on brain structure. However, we still have no sufficient knowledge to be able to represent the brain. Hence, it is doubted whether neural networks have the correct structure. Rule learning can claim some psychological backing while the stochastic belief learning model is based on currently available psychological knowledge. However, the stochastic belief learning model requires some knowledge on the potential beliefs of individuals that is often not available. In addition, it is more complex than fictitious play and the rule learning model.
Bayesian learning and least-squares learning are not supported by psychological knowledge. Two arguments can be put forward against Bayesian learning. First, people are not able to

do the calculations for a proper Bayesian updating. Second, people do not consider a large number of competing expectations or hypothesis at one time. Usually individuals have one specific expectation about reality. Psychologists argue that people tend to fix schemas and scripts very quickly even if little evidence has been collected (see Dörner 1999) and that they do not change these for more adequate alternatives if they lead to satisficing results (see Luchins 1942 and Anderson 2000). Similar to Bayesian learning, least-square learning does not fit the psychological evidence related to cognitive learning. People are simply either not able or not willing to do demanding calculations in most real situations as assumed in least-squares learning.

Genetic programming can be used to describe one individual or a population of agents (see, e.g. Edmunds 1999 and Chen & Yeh 2001). Presenting each agent by a population of formulas or programs means that we assume that different beliefs compete in the agent's brain and are simultaneously developed. This contradicts the above psychological finding that people usually only have one mental model at any one time. Presenting a population of agents by a population of formulas or programs implies that agents copy the beliefs of others. It has been stated above that it is difficult to study beliefs because they cannot be observed easily. Hence, it is unclear how agents can perfectly copy beliefs as is assumed in such an approach. In addition, cross-over operations are difficult to interpret. Why do agents exchange part of their beliefs instead of one agent convincing another of her beliefs? The advantage of genetic programming is that it allows the learning process to be very open with respect to the resulting beliefs.

Classifier systems have the interesting feature that they also model the development of a classification of situations. All other available learning models describe the learning process in one given situation. Thus, classifier systems focus on an element of cognitive learning that other learning models ignore: the development and change of schemas. However, this comes at some cost. Classifier systems do not represent beliefs in the same way the other belief learning models do. Instead, classifier systems define simple condition-action rules. They fail to accurately describe the learning of beliefs, which has been declared the central feature of belief learning here.

To sum up, we argue that more research on the modelling of belief learning is necessary. Given the current knowledge, we recommend the models of fictitious play and rule learning as simpler solutions and the stochastic belief learning model if more knowledge about beliefs is available. If the invention of new beliefs by individuals is an important feature of the processes under investigation, genetic programming could be also an option.

For the future we can hope that more empirical and experimental tests are conducted for the various learning models. This would help to develop a clearer picture of the conditions for different learning process to occur and the accurate ways to model them.

References

Anderson, C.M. & Camerer, C. (2000), Experience-Weighted Attraction Learning in Sender-Receiver Signaling Games, Economic Theory 16: 689-718.

Anderson, J.R. (2000), Cognitive Psychology and Its Implications (Worth Publishers, New

York).

Arifovic, J. (1994), Genetic Algorithm Learning and the Cobweb Model, Journal of Economic Dynamics and Control 18: 3-28.

Arifovic, J. & Ledyard, J. (2000), Scaling Up Learning Models in Public Good Games, Journal of Public Economic Theory 6: 203-238.

Arthur, W.B. (1991), Designing Economic Agents that Act Like Human Agents: A Behavioral Approach to Bounded Rationality, The American Economic Review 81: 353-359.

Bäck, T. (1996), Evolutionary Algorithms in Theory and Practice (Oxford University Press, New York).

Bandura, A. (1977), Social Learning Theorie (Prentice-Hall, Englewood Cliffs).

Bendor, J., Mookherjee, D. & Ray, D. (2001), Aspiration-based Reinforcement Learning in Repeated Interaction Games: An Overview, International Game Theory Review 3: 159-174.

Bergemann, D. & Välimäki, J. (1996), Learning and Strategic Pricing, Econometrica 64: 1125-1149.

Berninghaus, S.K. & Ehrhart, K.-M. (1998), Time Horizon and Equilibrium Selection in Tacit Coordination Games: Experimental Results, Journal of Economic Behavior and Organization 37: 231-248.

Beltrametti, L., Fiorentini, R., Marengo, L. & Tamborini, R. (1997), A Learning-to-Forcast Experiment on the Foreign Exchange Market with a Classifier System, Journal of Economic Dynamics & Control 21: 1543-1575.

Beltratti, A., Margarita, S. & Terna, P. (1996), Neural Networks for Economic and Financial Modelling (International Thomson Computer Press, London).

Beyer, H.-G. (2001), The Theory of Evolution Strategies (Springer, Berlin).

Biel, A. & Dahlstrand, U. (1997), Habits and the Establishment of Ecological Purchase Behavior, in: G. Guzmán, A. José & S. Sanz, eds., The XXII International Colloquium of Economic Psychology (Promolibro, Valencia) 367-381.

Börgers, T. (1996), On the Relevance of Learning and Evolution to Economic Theory, The Economic Journal 106: 1374-1385.

Börgers, T. & Sarin, R. (1997), Learning Through Reinforcement and Replicator Dynamics, Journal of Economic Theory 77: 1-16.

Börgers, T. & Sarin, R. (2000), Naive Reinforcement Learning with Endogenous Aspirations, International Economic Review 41: 921-950.

Bosch-Domenech, A. & Vriend, N.J. (2003), Imitation of Successful Behaviour in Cournot Markets, Economic Journal 113: 495-524.

Bray, M. (1982), Learning, Estimation, and the Stability of Rational Expectations, Journal of Economic Theory 26: 318-339.

Brenner, T. (1997), Reinforcement Learning in 2x2 Games and the Concept of Reinforcably Stable Strategies, Papers on Economics & Evolution #9703, Max Planck Institute, Jena.

Brenner, T. (1998), Can Evolutionary Algorithms Describe Learning Processes ?, Journal of Evolutionary Economics 8: 271-283.

Brenner, T. (1999), Modelling Learning in Economics (Edward Elgar, Cheltenham).

Brenner, T. (2001), Implications of Routine-based Learning for Decision Making, European Journal of Economic and Social Systems 15: 131-152.

Brenner, T. (2005), Stochastic Belief Learning and the Emergence of Cooperation, mimeo, Max Planck Institute, Jena.

Brenner, T. & Hennig-Schmidt, H. (2005), Belief Learning in the Prisoner's Dilemma Game - An Experimental Study of Cognitive Processes, mimeo, Max Planck Institute, Jena.

Brenner, T. & Murmann, J.P. (2003), The Use of Simulations in Developing Robust Knowledge about Causal Processes: Methodological Considerations and an Application to Industrial Evolution, Papers on Economics & Evolution #0303, Max Planck Institute, Jena.

Brenner, T. & Slembeck, T. (2001), Noisy Decision Makers - On Errors, Noise and Inconsistencies in Economic Behaviour, Papers on Economics & Evolution #0107, Max Planck Institute, Jena.

Brenner, T. & Vriend, N. (2005), On the Behavior of Proposers in Ultimatum Games, Journal of Economic Behavior and Organisation : forthcoming.

Brenner, T. & Witt, U. (2003), Melioration Learning in Games with Constant and Frequency-dependent Payoffs, Journal of Economic Behavior and Organization 50: 429-448.

Brown, G. W. (1951), Iterative Solution of Games by Fictitious Play, in: ??, Activity Analysis of Production and Allocation (John Wiley and Sons, ??) 374-376.

Bruner, J. S. (1973), Beyond the Information Given (Norton, New York).

Bullard, J. & Duffy, J. (1994), Learning in a Large Square Economy, Working Paper 94-013A, Federal Reserve Bank of St. Louis.

Bush, R. R. & Mosteller, F. (1955), Stochastic Models for Learning (John Wiley & Sons, New York).

Calderini, M. & Metcalfe, S. (1998), Compound Learning, Neural Nets and the Competitive Process, Economics of Innovation & New Technology 7: 271-302.

Camerer, C.F. & Ho, T. (1999), Experience-weighted Attraction Learning in Normal Form Games, Econometrica 67: 827-874.

Chen, S.-H., Duffy, J. & Yeh, C.-H. (2002), Equilibrium Selection via Adaptation: Using Genetic Programming to Model Learning in a Coordination Game, The Electronic Journal of Evolutionary Modeling and Economic Dynamics no. 1002: http://www.e-jemed.org/1002/index.php.

Chen, S.-H. & Yeh, C.-H. (2001), Evolving Traders and the Business School with Genetic Programming: A New Architecture of the Agent-Based Artificial Stock Market, Journal of Economic Dynamics & Control 25: 363-393.

Cheung, Y.-W. & Friedman, D. (1997), Individual Learning in Normal Form Games: Some Laboratory Results, Games & Economic Behavior 19: 46-76.

Crawford, V. P. (1995), Adaptive Dynamics in Coordination Games, Econometrica 63: 103-143.

Cross, J. G. (1973), A Stochastic Learning Model of Economic Behavior, Quarterly Journal of Economics 87: 239-266.

Cross, J. G. (1983), A Theory of Adaptive Economic Behavior (Cambridge University Press, Cambridge).

Dawid, H. (1996), Adaptive Learning by Genetic Algorithms (Springer, Lecture Notes in Economics and Mathematical Systems 441, Berlin).

Dawid, H. (1997), Learning of Equilibria by a Population with Minimal Information, Journal of Economic Behavior & Organization 32: 1-18.

Day, R. H. (1967), Profits, Learning and the Convergence of Satisficing to Marginalism, Quarterly Journal of Economics 81: 302-311.

Day, R. H. & Tinney, E. H. (1968), How To Co-operate in Business without Really Trying: A Learning Model of Decentralized Decision Making, Journal of Political Economy 74: 583-600.

Denzau, A. & North, D. C. (1994), Shared Mental Models. Ideologies and Institutions, Kyklos 47: 3-31.

Dixon, H. D. (2000), Keeping Up With the Joneses: Competition and the Evolution of Collusion, Journal of Economic Behavior & Organization 43: 223-238.

Dörner, D. (1999), Bauplan für eine Seele (Rowohlt, Hamburg).

Duffy, J. & Feltovich, N. (1999), Does Observation of Others Affect Learning in Strategic Environments? An Experimental Study, International Journal of Game Theory 28: 131-152.

Duffy, J. (2005), Agent-Based Models and Human-Subject Experiments, in this handbook.

Easley, D. & Kiefer, N. M. (1988), Controlling a Stochastic Process with Unknown Parameters, Econometrica 56: 1045-1064.

Eichberger, J., Haller, H. & Milne, F. (1993), Naive Bayesian Learning in 2x2 Matrix Games, Journal of Economic Behavior and Organization 22: 69-90.

Eigen, M. (1971), The Self-Organisation of Matter and the Evolution of Biological Macromolecules, Naturwissenschaften 58: 465-523.

Edmunds, B. (1999), Modelling Bounded Rationality in Agent-Based Simulations Using the Evolution of Mental Models, in: T. Brenner, ed., Computational Techniques for Modelling Learning in Economics (Edward Elgar, Cheltenham) 305-332.

Ellison, G. (1993), Learning, Local Interaction, and Coordination, Econometrica 61: 1047-1071.

Erev, I., Bereby-Meyer, Y. & Roth, A.E. (1999), The Effect of Adding a Constant to All Payoffs: Experimental Investigation, and Implications for Reinforcement Learning Models, Journal of Economic Behaviour and Organization 39: 111-128.

Erev, I. & Roth, A.E. (1998), Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria, The American Economic Review 88: 848-881.

Eshel, I., Samuelson, L. & Shaked, A. (1998), Altruists, Egoists and Hooligans in a Local Interaction Model, American Economic Review 88: 157-179.

Estes, W. K. (1950), Toward a Statistical Theory of Learning, Psychological Review 57: 94-107.

Feltovich, N. (2000), Reinforcement-based vs. Belief-based Learning Models in Experimental Asymmetric-information Games, Econometrica 68: 605-641.

Festinger, L. (1942), Wish, Expectation and Group Standards as Factors Influencing the Level of Aspiration, Journal of abnormal social Psychology 37: 184-200.

Fudenberg, D. & Levine, D.K. (1998), The Theory of Learning in Games (MIT Press, Cambridge).

Gigerenzer, G. & Selten, R. (eds.) (2001), Bounded Rationality: The Adaptive Toolbox (MIT Press, Cambridge).

Gilboa, I. & Schmeidler, D. (1995), Case-based Decision Theory, Quarterly Journal of Economics 110: 269-288.

Gilboa, I. & Schmeidler, D. (2001), Reaction to Price Changes and Aspiration Level Adjustment, Review of Economic Design 6: 215-223.

Gößling, T. (1996), Entscheidung in Modellen (Hänsel-Hohenhausen Verlag der Deutschen Hochschulschriften, Egelsbach).

Goldvarg, E. & Johnson-Laird, P.N. (2001), Naive Causality: A Mental Model Theory of Causal Meaning and Reasoning, Cognitive Science 25: 565-610.

Grosskopf, B. (2003), Reinforcement and Directional Learning in the Ultimatum Game with Responder Competition, Experimental Economics 6: 141-158.

Hallpike, C. R. (1986), The Principles of Social Evolution (Clarendon Press, Oxford).

Hebb, D.O. (2002), The Organization of Behavior. A Neuropsychological Theory (Lawrence Erlbaum Associates, Mahwa).

Hegselmann, R. (1996), Social Dilemmas in Lineland and Flatland, in: W. B. G. Liebrand & D. M. Messick, eds., Frontiers in Social Dilemmas Research (Springer, Berlin) 337-361.

Heinemann, M. (2000), Adaptive Learning of Rational Expectations Using Neural Networks, Journal of Economic Dynamics & Control 24: 1007-1026.

Helbing, D. (1995), Quantitative Sociodynamics, Stochastic Methods and Models of Social Interaction Processes (Kluwer Academic, Boston).

Herrnstein, R. J. (1970), On the Law of Effect, Journal of the Experimental Analysis of Behavior 13: 243-266.

Herrnstein, R. J. & Prelec, D. (1991), Melioration: A Theory of Distributed Choice, Journal of Economic Perspectives 5: 137-156.

Hofbauer, J. & Sigmund, K. (1984), Evolutionstheorie und dynamische Systeme (Paul Parey, Berlin).

Holland, J. H. (1975), Adaption in Natural and Artificial Systems (University of Michigan Press, Ann Arbor).

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

Holland, J. H. & Miller, J. H. (1991), Artifical Adaptive Agents in Economic Theory, American Economic Review, Papers and Proceedings 81: 365-370.

Huck, S., Normann, H.-T. & Oechssler, J. (2002), Stability of the Cournot Process - Experimental Evidence, Internation Journal of Game Theory 31: 123-136.

Johnson-Laird, P. N. (1983), Mental Models (Harvard University Press, Cambridge).

Jordan, J. S. (1991), Bayesian Learning in Normal Form Games, Games and Economic Behavior 3: 60-81.

Jordan, J. S. (1995), Bayesian Learning in Repeated Games, Games and Economic Behavior 9: 8-20.

Kalai, E. & Lehrer, E. (1993), Rational Learning Leads to Nash Equilibrium, Econometrica 61: 1019-1045.

Kahneman, D. & Tversky, A. (1979), Prospect Theory: An Analysis of Decision Under Risk, Econometrica 47: 263-291.

Kahneman, D. (2003), Maps of Bounded Rationality: Psychology for Behavioral Economics, The American Economic Review 93: 1449-1475.

Kandel, E. R. & Schwartz, J. H. (1982), Molecular Biology of Learning: Modulation of Transmitter Release, Science : 433-443.

Kandori, M., Mailath, G.J. & Rob, R. (1993), Learning, Mutation, and Long Run Equilibria in Games, Econometrica 61: 29-56.

Karandikar, R., Mookherjee, D., Ray, D. & Vega-Redondo, F. (1998), Evolving Aspirations and Cooperation, Economic Theory 80: 292-331.

Kirchkamp, O. (2000), Spatial Evolution of Automata in the Prisoner's Dilemma, Journal of Economic Behavior and Organization, 43.239-262

Kubon-Gilke, G. (1997), Verhaltensbindung und die Evolution ökonomischer Institutionen (Metropolis-Verlag, Marburg).

Lant, T. K. (1992), Aspiration Level Adaption: An Empirical Exploration, Management Science 38: 623-644.

Luchins, A.S. (1942), Mechanization in Problem-solving, Psychological Monographs 54: No. 248.

Mackintosh, N.J. (2003), Pavlov and Associationism, The Spanish Journal of Psychology 6: 177-184.

Marcet, A. & Sargent, T. J. (1989), Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models, Journal of Economic Theory 48: 337-368.

Maynard Smith, J. (1982), Evolution and the Theory of Games (Cambridge University Press, Cambridge).

Mezias, S.J. (1988), Aspiration Level Effects: An Empirical Investigation, Journal of Economic Behavior and Organization 10: 389-400.

Mezias, S., Chen, Y. & Murphy, P. (2002), Aspiration-Level Adaptation in an American Financial Service Organization: A Field Study, Management Science 48: 1285-1300.

Morales, A.J. (2002), Absolutely Expedient Imitative Behavior, International Journal of Game Theory 31: 475-492.

Nagel, R. (1995), Unraveling in Guessing Games: An Experimental Study, American Economic Review 85: 1313-1326.

Nash, J. F. (1950), Equilibrium Points in n-Person Games, Proceedings of the National Academy of Sciences (US) 36: 48-49.

Nowak, M. A. & May, R. M. (1993), The Spatial Dilemmas of Evolution, International Journal of Bifurcation and Chaos 3: 35-78.

Nyarko, Yaw & Schotter, Andrew (2002), An Experimental Study of Belief Learning Using Elicited Beliefs, Econometrica 70: 971-1005.

Osborne, M.J. & Rubinstein, A. (1995), A Course in Game Theory (MIT Press, Cambridge).

Palomino, F. & Vega-Redondo, F. (1999), Convergence of Aspiration and (Partial) Cooperation in the Prisoner's Dilemma, International Journal of Game Theory 28: 465-488.

Pazgal, A. (1997), Satisfycing Leads to Cooperation in Mutual Interests Games, International Journal of Game Theory 26: 439-453.

Piaget, J. (1976), Die Äquilibration der kognitiven Strukturen (Klett, Stuttgart).

Ramstad, Y. (1994), On the Nature of Economic Evolution, in: L. Magnusson, ed., Evolutionary and Neo-Schumpeterian Approaches to Economics (Kluwer Academic Publishers, Boston) .

Rapoport, A., Seale, D. & Winter, E. (2000), An Experimental Study of Coordination and Learning in Iterated Two-Market Entry Games, Economic Theory 16: 661-687.

Rechenberg, I. (1973), Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution (Frommann-Holzboog Verlag, Stuttgart).

Rescorla, R. A. & Wagner, A. R. (1972), A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Non-reinforcement, in: A. H. Black & W. F. Prokasy, eds., Classical Conditioning II (Prentice Hall, Englewood Cliffs) ..

Robinson, J. (1951), An Iterative Method of Solving a Game, Annals of Mathematics 54: 296-301.

Roth, A.E. & Erev, I. (1995), Learning in Extensive Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Run, Games and Economic Behavior 6: 164-212.

Rumiati, R.I. & Bekkering, H. (2003), To Imitate or Not to Imitate? How the Brain Can Do It, That Is the Question!, Brain and Cognition 53: 479-482.

Samuelson, L. (1994), Stochastic Stability in Games with Alternative Best Replies, Journal of Economic Theory 64: 35-65.

Sarin, R. & Vahid, F. (1999), Payoff Assessments without Probabilities: A Simple Dynamic Model of Choice, Games and Economic Behavior 28: 294-309.

Schlag, K.H. (1998), Why Imitate, and if so, How? A Bounded Rational Approach to Multi-armed Bandits, Journal of Economic Theory 78: 130-156.

Schwefel, H.-P. (1995), Evolution and Optimum Seeking (Wiley & Sons, New York).

Selten, R. (1997), Features of Experimentally Observed Bounded Rationality, Discussion Paper No. B-421, University of Bonn.

Selten, R. & Stoecker, R. (1986), End Behavior in Sequences of Finite Prisoner's Dilemma Supergames, Journal of Economic Behavior and Organization 7: 47-70.

Shapley, L. S. (1964), Some Topics in Two-Person Games, in: M. Dresher, L. Shapley & A. Tucker, eds., Advances in Game Theory (Annals of Mathematic Studies No. 52, Princeton) 1-28.

Simon, H. A. (1957), Administrative Behavior (The Macmillan Company, New York).

Simon, H. A. (1987), Satisficing, The New Palgrave Dictionary of Economics Vol. 4, Macmillan Press, London: 243-245.

Skinner, B. F. (1938), The Behavior of Organisms (Appleton, New York).

Stahl, D.O. (1998), Is Step-j Thinking an Arbitrary Modelling Restriction or a Fact of Human Nature?, Journal of Economic Behavior and Organisation 37: 33-51.

Stahl, D.O. (2000), Rule Learning in Symmetric Normal-Form Games: Theory and Evidence, Games and Economic Behavior 32: 105-138.

Stahl, D.O. & Haruvy, E. (2002), Aspiration-Based and Reciprocity-Based Rules of Learning Dynamics for Symmetric Normal-Form Games, Journal of Mathematical Psychology 46: 531-553.

Stahl, D.O. & Wilson, P.W. (1994), Experimental Evidence on Player's Models of Other Players, Journal of Economic Behavior and Organization 25: 309-327.

Tang, F.-F. (2003), A Comparative Study on Learning in a Normal Form Game Experiment, Journal of Economic Behavior and Organization 50: 385-390.

Tesfatsion, L. (2001), Introduction to the Special Issue on Agent-Based Computational Economics, Journal of Economic Dynamics & Control 25: 281-293.

Thibaut, J. W. & Kelley, H. H. (1959), The Social Psychology of Groups (Wiley, New York).

Thorndike, E. L. (1932), The Fundamentals of Learning (AMS Press, New York).

Vega-Redondo, F. (1997), The Evolution of Walrasian Behavior, Econometrica 65: 375-384.

Vaughan, W. & Herrnstein, R. J. (1987), Stability, Melioration, and Natural Selection, in: L. Green, J. H. Kagel, eds., Advances in Behavioral Economics, vol. 1 (Ablex, Norwood) .

Vriend, N.J. (2000), An Illustration of the Essential Difference Between Individual and Social Learning, and Its Consequences for Computational Analysis, Journal of Economic Dynamics & Control 24: 1-19.

Werker, C. & Brenner, T. (2004), Empirical Calibration of Simulation Models, Papers on Economics & Evolution #0410, Max Planck Institute, Jena.

Witt, U. (1986), Firms' Market Behavior under Imperfect Information and Economic Natural Selection, Journal of Economic Behavior and Organization 7: 265-290.

Witt, U. (1991), Evolutionary Economics – an Interpretative Survey, Papers on Economics & Evolution # 9104, European Study Group for Evolutionary Economics, Freiburg.

Witt, U. (1996), Bounded Rationality, Social Learning, and Viable Moral Conduct in a Prisoners' Dilemma, in: M. Perlman & E. Helmstädter, eds., Behavioral Norms, Technological Progress and Economic Dynamics: Studies in Schumpeterian Economics (Michigan University Press, Ann Arbor) 33-49.

Witt, U. (2001), Learning to Consume - A Theory of Wants and the Growth of Demand, Journal of Evolutionary Economics 11: 23-36.

Yin, G. & Zhu, Y. M. (1990), On H-Valued Robbins-Monro Processes, Journal of Multivariate Analysis 34: 116-140.

Young, P. (1993), The Evolution of Conventions, Econometrica 61: 57-84.

Zellner, A. (1971), An Introduction to Bayesian Inference in Econometrics (John Wiley, New York).

Table 1:
Classification according to the source of the learning models and according to the classification developed below.

<table>
<tr><th></th><th>non-conscious
learning</th><th>routine-based
learning</th><th>belief
learning</th></tr>
<tr><td>psychology-
based
models</td><td>Bush-Mosteller
model,
parameterised
learning automaton</td><td>satisficing, melioration,
imitation,
Roth-Erev model,
VID model</td><td>stochastic
belief
learning,
rule learning</td></tr>
<tr><td>rationality-
based models</td><td></td><td></td><td>Bayesian learning,
least-squares
learning</td></tr>
<tr><td>adaptive
models</td><td></td><td>learning
direction theory</td><td></td></tr>
<tr><td rowspan="2">belief
learning models</td><td></td><td></td><td>fictitious play</td></tr>
<tr><td colspan="3">EWA model</td></tr>
<tr><td>models
from AI
and
biology</td><td></td><td>evolutionary algorithms,
replicator dynamics,
selection-mutation
equation</td><td>genetic
programming,
classifier systems
neural networks</td></tr>
</table>

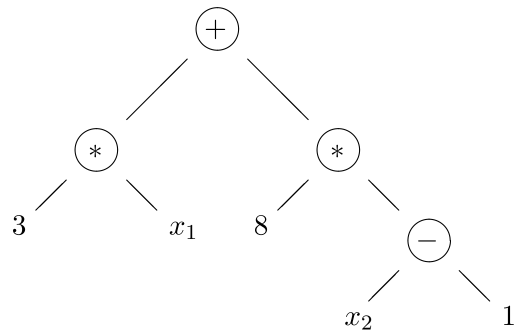

Figure 1: Coding in genetic programming, an example.

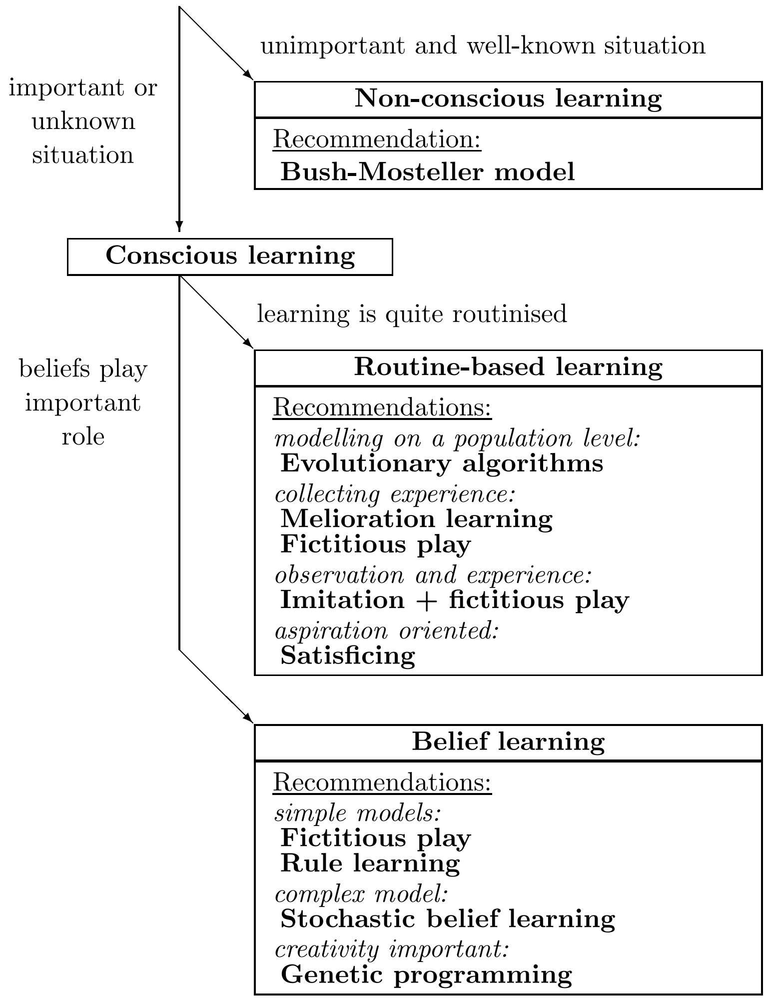

Figure 2: Steps to choose an accurate learning model for representing economic behaviour.