Reasoning in the Absence of Goals

Mary Lou Maher
University of Maryland
mlmaher@umd.edu

Kathryn E Merrick
University of New South Wales
k.merrick@adfa.edu.au

Benjamin Graham
University of New South Wales
benjamin.graham@student.adfa.edu.au

Abstract

In creative industries such as design and research it is common to reason about 'problem-finding' before tasks or goals can be established. Problem-finding may also continue throughout the problem-solving process, so achieving goals may be an ongoing process of discovery as well as iterative improvement and refinement. This paper considers the design of cognitive systems with complementary processes for both problem-finding and problem-solving. We review a range of approaches that may complement goal-directed reasoning when an artificial system does not or cannot know precisely what it is looking for. We argue that there is a spectrum of approaches that can be used for reasoning in the absence of goals, which make progressively weaker assumptions about the definition and presence goals, and that goal-oriented behavior can be an intermediate result of problem-finding, rather than as a starting point for problem-solving. We demonstrate one such approach based on implicit motives and incentives.

Introduction

AI approaches to cognitive systems assume that explicit representations of goals, rewards and tasks are integral and provide a focus of attention. Such approaches imply that we are modeling cognitive systems with the assumption that goals are the starting point for reasoning; that reasoning cannot start without goals; and that reasoning ends when there are no goals. In contrast, this paper characterizes reasoning so that goals become flexible intermediate structures or implied structures, rather than a predefined and fixed starting point. By using concepts such as incentive, novelty, difficulty, complexity, curiosity and surprise, cognitively inspired AI models that mimic human behavior in scenarios such as exploratory research and lifelong, self-directed learning are possible. The models can be applied to any domain in which 'problem-finding' needs to occur during problem solving.

The remainder of this section overviews existing approaches to goal-oriented behavior in cognitive systems. The next section examines a number of complementary approaches that may work in conjunction with goal-directed reasoning, including incentives, motivation, surprise, novelty and curiosity. We classify these approaches along a spectrum that makes progressively weaker assumptions about the definition and presence of goals. One approach at the weaker end of this spectrum is demonstrated based on implicit motives and incentives. We show that goal-oriented behavior can be understood as an intermediate result of problem-finding, rather than as a starting point for problem solving and that goal-oriented behavior can be an emergent property that does not depend on explicit definition of domain-specific goals.

Copyright © 2011, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Goal-Oriented Behavior in Cognitive Systems

Langley et al. (2008) survey cognitive architectures and layout nine functional capabilities that are required of a cognitive architecture. While each of these are described in terms of functionality without reference to specific architectures or implementations, almost all assume that goals and tasks are inherent in the description of the functions. This can be seen in the way that Langley et al (2008) describe four examples of cognitive architectures:

- Soar: "**All tasks** in Soar are formulated as attempts **to achieve goals**."
- ACT-R: "ACT-R 6 is organized into a set of modules, each of which processes a different type of information. These include sensory modules for visual processing, motor modules for action, an **intentional module for goals**, and a declarative module for long-term declarative knowledge."
- ICARUS: "ICARUS … stores two distinct forms of knowledge. Concepts describe classes of environmental situations in terms of other concepts and percepts, whereas **skills specify how to achieve goals** by decomposing them into ordered subgoals."
- PRODIGY: "On each cycle, PRODIGY uses its control rules to select an operator, binding set, state, or **goal,** to reject them out of hand, or to prefer some over others. In the absence of such control knowledge, the architecture makes choices at random and pursues depth-first means-ends search with backtracking"

There are multiple models for goals – including goal lifecycles and type taxonomies (Braubach et al., 2005) – and processes for solving goals – including machine learning (Nilsson, 1996), planning and rule-based agents (Russell and Norvig, 1995). Braubach et al., (2005) define a lifecycle for goals in which goals transition from new to adopted and finished.

Braubach et al., (2005) also divide goals into a number of types. Approach goals, for example, define states for which an agent should minimize the difference between its current state and the goal state. In contrast, avoidance goals define states for which an agent should maximize the difference between its current state and the goal state. Achievement goals define changes or events that the agent should cause to occur. Maintenance goals define properties that the agent should hold constant. Other types of goals include optimization, test, query and cease goals.

Dignum and Conte (1998) state that truly autonomous, intelligent agents must be capable of creating new goals as well as dropping goals as conditions change. They distinguish between abstract, high-level goals and concrete, achievable goals. They describe goal formation as a process of deriving concrete, achievable goals – such as 'driving at the speed limit' – from high level, abstract goals – such as 'being good'.

Foner and Maes (1994) develop an agent model of unsupervised learning that can self-determine what facts it should pay attention to as a way of modeling focus of attention. Foner and Maes (1994) distinguish between goal-driven and world-driven focus of attention. In their model, the agent can determine what sensory data to learn from based on strategies that are derived from world-driven goals, such as what has changed recently and what new data is spatially close. These are domain independent strategies that can reduce the number of possible goals an agent can pursue at any given time.

In general, however, there has been less work on how to represent the high-level, abstract goals or world-driven goals that cause new, concrete goals to emerge. The concept of an abstract goal is difficult to formalize because of the difficulty of representing high-level objectives such as "being good" or "being creative". A number of alternative approaches use models of motivation to take the place of abstract learning goals (Merrick and Maher, 2009; Singh et al., 2005; Kaplan and Oudeyer, 2003; Schmidhuber, 1991). Computational models of motivation have also been proposed as an approach to embedding implicit motives in artificial agents to create agents with different preferences for certain kinds of activities (Merrick and Shafi, 2011). In a different approach, Barnes and Oudeyer (2010) presented a framework for 'maturationally-constrained self-adaptive goal generation' in which an intrinsic motivation module progressively releases constraints on the learning system. This permits the learning system to explore progressively more widely, through the introduction of new goals.

Other work has studied the role of emotion and other cognitive moderators in artificial systems (Mariner and Laird, 2008). Models of emotion act as modifiers to an agent's goal-oriented behavior or provide abstract goals that can be mapped onto concrete goals.

This paper proposes that reasoning can include reasoning before goals are defined, usually based on the current state of the artificial agent and the state of the world. These approaches are consistent with current cognitive systems because they ultimately lead to goal-oriented behavior, but they complement most cognitive systems because they do not assume that goals are predefined. The next section describes models that fall along a spectrum that make progressively weaker assumptions about the definition and presence of goals.

Models that Complement Goal-Directed Reasoning

In creative domains such as design and research, the ill-defined nature of tasks suggests a distinction between search and exploration. Maher et al (1996) characterize the difference between search and exploration by the input and output of these processes as illustrated in Figure 1. A typical search process generates a solution as its output with a well-defined problem (or goal) as its input. However, an exploration process derives a problem and the corresponding solution from an ill-defined problem. Maher et al (1996) take this idea further to propose a co-evolutionary model of reasoning about the problem space and the solutions space in which goals are expressed as requirements that change in response to the evolutionary search of the solution space.

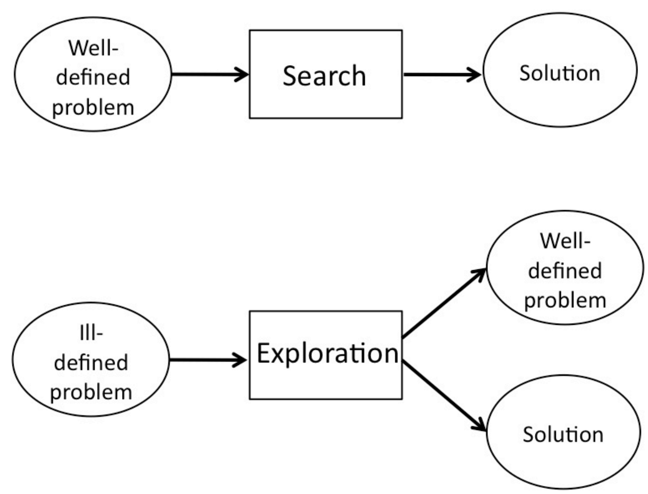

Figure 1. Input and output of search and exploration (Maher et al. 1996).

If we consider the internal state of an agent to include its goals, then the absence of goals remains a valid state. In many autonomous systems the absence of goals implies idle time, but we envisage that a cognitive system can continue to monitor its environment to discover and pursue self-generated goals that extend or improve its knowledge base or skill set, during this so-called idle time. This type of activity will cause changes in the agent's internal and external environment and create a feedback loop that fosters continuous adaptation.

While ultimately the processes in cognitive systems are organized with the assumption that behavior is goal-

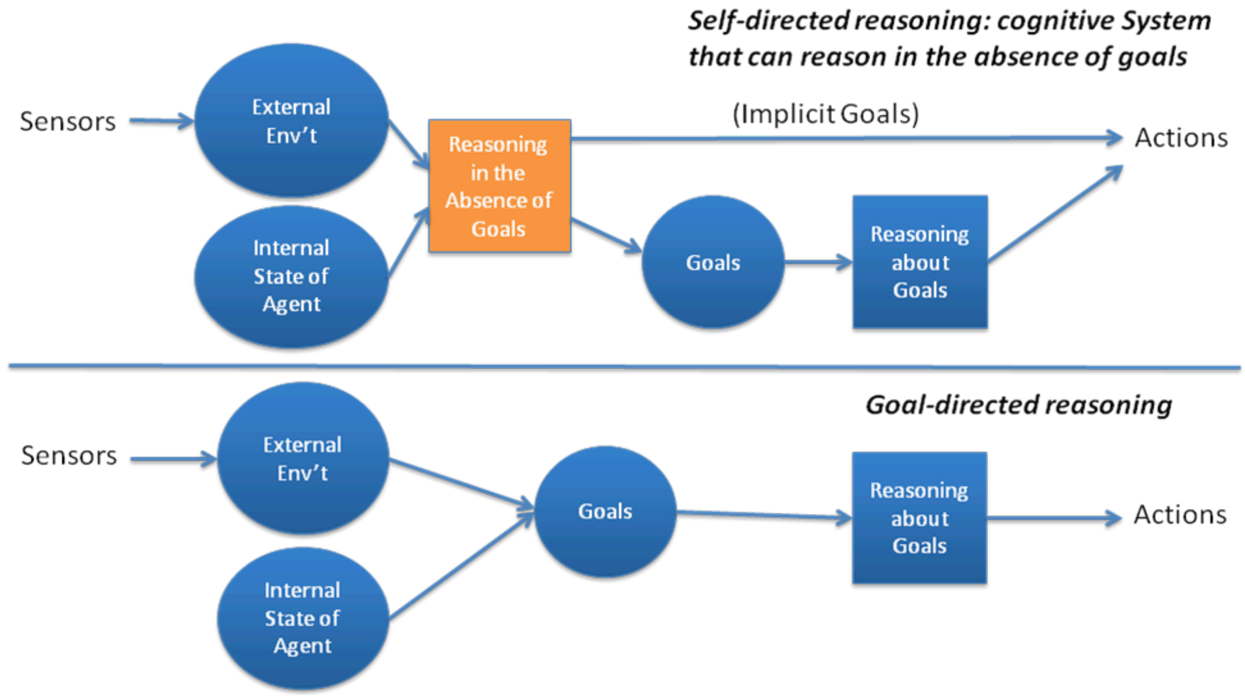

Figure 2: The role of reasoning in the absence of goals in a self-directed cognitive system

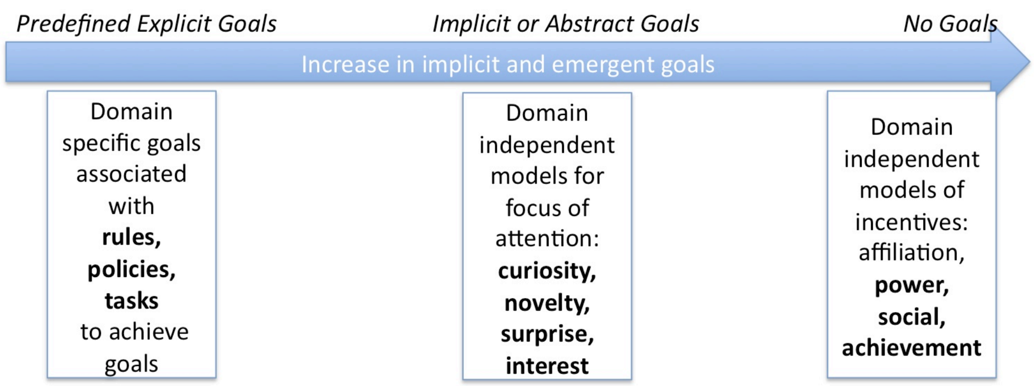

Figure 3. Models for reasoning with and without goals.

directed, we propose that self-directed cognitive systems include the ability to represent, generate, and reason about what Dignum and Conte (1998) call abstract goals. Rather than cast this capability in terms of goals and tasks, however, we identify cognitive models that complement goal-directed reasoning. This is illustrated in Figure 2, where reasoning in the absence of goals can lead to action without an explicit representation of goals or it can lead to the definition of new goals for goal-directed reasoning.

In self-directed reasoning, goals are flexible intermediate structures or implied structures, rather than a predefined and fixed starting point for reasoning. Figure 3 shows a spectrum of reasoning starting with the traditional goal-directed reasoning that includes domain specific goals and models for achieving them, through an intermediate type of reasoning in which goals are implied and may be emergent properties of reasoning, to reasoning without goals in which incentives provide guidance for reasoning about actions. Goals can be flexible, intermediate structures, rather than fixed starting points. This will permit continuous learning and adaptation to unexpected data or events, or changes in needs, beliefs or desires to become an integral concept in cognitive architectures. This will also recast cognitive systems from strictly goal-directed behaviors to include creative and exploratory behaviors.

In this section we begin by looking at models that represent abstract goals as in terms of reward, such as novelty, curiosity, interest and surprise. We then consider a set of weaker models that represent only a preference for certain types of abstract goals and not predefined domain specific goals. We then take the latter of these (implicit motives and incentives) and present a model that demonstrates how goals can be thought of as an emergent output of motivation, rather than an explicitly represented starting point for reasoning.

Novelty, Interest and Curiosity

Novelty, interest and curiosity fall in a class of models that allow an agent to respond to and learn from changes in the environment. There are many accounts of measuring novelty using computational approaches. Marsland et al. (2000) used Stanley's (1976) model of habituation to implement a real-time novelty detector for mobile robots. Like the Kohonen (1993) Novelty Filter, the real-time novelty detector uses a Self-Organising Map (SOM) as the basis for the detection of novelty. Habituation and recovery extends a novelty filter with the ability to forget.

Models of interest provide a basis for determining if a novel event or state is worth a focus of attention. Curiosity is when something of interest can distract the process from its current focus of attention. Saunders and Gero (2001) drew on the work of Berlyne (1960) and Marsland et al (2000) to develop computational models of curiosity and interest based on novelty. They used a real-time novelty detector to implement novelty. Saunders and Gero (2004) model interest using sigmoid functions to represent positive reward for the discovery of novel stimuli and negative reward for the discovery of highly novel stimuli. The resulting computational models of novelty and interest are used in a range of applications including curious agents.

Merrick and Maher (2009) present models of motivated reinforcement learning agents that use novelty and curiosity as models of intrinsic motivation. These agents exhibit a kind of world-driven (rather than goal-driven) behavior. The agents have an experience trajectory $Y_{(t)}$ that models all states $S_{(t)}$, changes in states (events) $E_{(t)}$, actions that have been encountered/experienced by the agent:

$$Y_{(t)} = S_{(1)}, E_{(1)}, A_{(1)}, S_{(2)}, E_{(2)}, A_{(2)}, \ldots , S_{(t)}, E_{(t)}, A_{(t)}$$

A dynamic motivated reward signal $R_{\mathrm{m}(t)}$ is computed as a function of novelty and interest. Their model of interest, based on the experience trajectory, is a modified version of the Saunders and Gero interest function and is based on the Wundt curve shown in Figure 4.

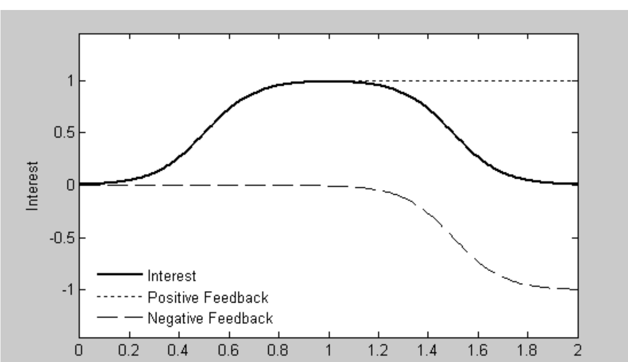

Figure 4 The Wundt curve is the difference between positive and negative feedback functions. It peaks at a moderate degree of novelty (Merrick and Maher, 2009).

This curiosity-based reward signal directs the agent to focus its learning on achieving specific situations at different times, but does not have an explicit representation of tasks or goals. Other motivation functions studied by Merrick and Maher (2009) within this framework include functions for competency and combined competency and curiosity. Experimental studies of curious agents in dynamic environments demonstrated highly adaptive behaviors through the ability to learn a high variety of simple and complex behaviors (see Merrick and Maher, 2009 for experimental results).

Surprise

Surprise occurs when an unexpected event occurs. While surprise and novelty are similar, something may be novel, but necessarily surprising because it is the next expected change. Horvitz et al (2005) and Itti and Baldi (2004) have developed probabilistic models for finding surprising events in data. Ranasinghe and Shen (2008) have developed a model of surprise for reinforcement learning for developmental robots.

The Horvitz et al (2005) model of surprise is used in traffic forecasting. They generated a set of probabilistic dependencies among a set of random variables, for example linking weather to traffic status. They assume a user model that states that when an event occurs that has less than 2% probability of occurring, it is marked as surprising. Surprising events in the past are collected in a case library of surprises. This provides the data for forecasting surprises based on current traffic conditions.The Itti and Baldi (2004) model of surprise is developed for observing surprising features in image data using a priori and posterior probabilities. Given a user dependent model M of some data, there is a $P(M)$ describing the probability distribution. $\mathrm{P}(M|D)$ is the probability distribution after the data is added, using Bayesian probability. Surprise is modeled as the distance d between the prior, $\mathrm{P}(M)$, and posterior $\mathrm{P}(M|D)$ probabilities.

The Ranasinghe and Shen (2008) model of surprise is used as a reward in a model they call surprise-based learning for developmental robots. In this model, surprise is used to set goals for learning in an unknown environment. The world is modeled as a set of rules, where each rule has the form: *Condition* → *Action* → *Predictions.* A condition is modeled as: *Feature* → *Operator* → *Value.* For example, a condition can be feature1 > value1 where "greater than" is the operator. A prediction is modeled as *Feature* → *Operator.* For example, a prediction can be "feature1 >" where it is expected that feature1 will increase after the action is performed. The comparison operators provided for surprise analysis include operators to detect the presence (%) or absence (~) of a feature, and the change in the size of a feature (<, <=, =, >=, >). If an observed feature does not match the prediction for the feature, for example, the feature was expected to increase and it decreased, then the system recognizes surprise and sets that state as a reward for learning.

The three models provide different approaches to modeling surprise based on the needs of the context in which they are developed. The Horvitz et al (2005) model determines that an event in the past is surprising, and then for a collection of surprising events is used to predict future surprising events. In the Itti and Baldi (2004) model, the new data is assimilated into the probability distribution, so something is surprising the first time it is introduced. The Ranasinghe and Shen (2008) model does not use probabilities and instead finds the first unexpected feature based on predictions of the direction in which the values of features will change and sets a reward to learn about that situation.

Incentives, Motives, and Motivation

In motivational psychology, incentive is defined as a situational characteristic associated with possible satisfaction of a motive (Heckhausen and Heckhausen, 2008). Incentives can be internal or external. Examples of internal incentives that depend on an individual's experiences include the novelty, difficulty or complexity of a situation. Examples of external incentives include money or other kinds of external 'payoff'. Associations between incentive and motivation can be learned, but there are also certain associations between incentives and motivation that have been found to be common across individuals. These include the associations between:

- Task difficulty and achievement motivation
- Risk and power motivation
- Risk and affiliation motivation
- Novelty and curiosity

Suppose we represent a situation encountered by an agent at time t as $S_{(t)}$. Then the incentives associated with a situation can be represented as $I_{(t)} = (i_1, i_2, i_3\ldots)$. Each value i_n represents a different incentive. For example, i_1 may describe the novelty of $S_{(t)}$, i_2 may describe the complexity of $S_{(t)}$, i_3 may describe risk and so on.

Internal incentive values such as novelty, difficulty and complexity can be computed by an agent while it is reasoning about its environment using computational models such as novelty-detectors (Marsland et al., 2000) or achievement based on error calculations on learned policies (Merrick and Maher, 2009). This means that the incentives associated with a situation will change based on the agent's experiences. External incentive values are interpreted from the current state of the environment $S_{(t)}$. These values will change based on changes in the environment. Both types of incentive have the possibility of satisfying the agent's motive.

Implicit motives are innate preferences for certain kinds of incentives. Because different individuals have different implicit motives, they will interpret the same situation incentives differently. For example, individuals with strong achievement motivation favor moderate difficulty. Likewise, high curiosity is associated with moderate novelty. Individuals with strong power motivation favor high risk. In contrast, individuals with strong affiliation motivation avoid situations with high risk. We can represent different motives M_1, M_2, $M_3\ldots$ as a function of incentive $M_{m(t)} = M_m(I_{(t)})$. These scalar motivation values $M_{m(t)}$ can be used in isolation, for example as a reward signal in learning, or combined. For example, they can be summed to give a resultant motivational tendency based on a complex motive profile of multiple motives (Merrick and Shafi, 2011).

$$T_{\mathrm{res(t)}} = M_{1(\mathrm{t})} + M_{2(\mathrm{t})} + M_{3(\mathrm{t})} + \ldots$$

The resultant value can then be used by the agent to identify the most highly motivating situations and act, learn to act or plan to act to achieve those situations. This action, learning or planning may involve formation of explicit concrete goal structures, but this is not strictly necessary. The demonstration in Section 4, for example, uses an agent architecture without an explicit goal representation, as do self-motivated learning algorithms such as motivated reinforcement learning.

In summary, incentives and motives are values derived from the synergy of the external environment and the internal state and preferences of the agent that provide a basis for deciding what to do next. Algorithm 2 is a continuous learning, planning, or action selection algorithm that uses a model of incentives and motives to determine what to do next.

```
Repeat:
        Sense state S(t)
        Compute incentives (internal & external) I(t) = (i1, i2, i3…).
        Compute Mm and Tres(t) = M1 + M2+ …
        If S(t) is highly motivating then
                Act, learn, plan, create a goal etc to achieve S(t)
        Else
                Act, learn, plan, create a goal etc to avoid S(t)
        End if
End
```

Algorithm 1. An algorithm for generating action or concrete goals for action from implicit motives.

Established motive profiles from psychology provide a starting point for designing different types of artificial agents. For example McClelland (1975) describe a leadership motive profile that combines high achievement and power motivation and results in individuals that emerge as leaders in a range of situations. Other profiles that have been studied include the power profile of high power motivation with low achievement and affiliation motivation. In artificial agents, this profile generates a kind of behavior or can be used to create abstract goals such as 'being aggressive' or being a 'risk-taker'. We demonstrate four motive profiles in the next section and show how they can be used to influence action and adaptation in artificial agents, in the absence of concrete goals.

Demo: Agents Playing Mixed-Motive Games

Mixed-motive games offer a simple representation of certain kinds of strategic interactions between individuals (players). In these games, players choose among a number of possible actions. The resulting payoff to each player is determined by the actions of both players. The best payoff is only obtained if both players act differently. This can be thought of as both players acting with different motives because the situation (game) itself is identical for both players. Mixed motive games provide a good environment for experimenting with and comparing alternatives to goal-directed behaviour, because they define a number of possible outcomes, but do not predefine goals. Rather, players need to dynamically determine which actions to select in a changing environment that is dependent on the actions of the other player(s).

Here we describe the effect of different players' motive profiles in a well known two-player game of 'Leader'. The Leader game models strategic interactions in which two players must decide whether to concede (*C*) or drive (*D*) with the outcomes being shown in the matrix as numbers. The numbers in Table 1 show a comparison of which outcomes are more desirable, with 4 being the most desirable payoff, 3 being the second best etc. The payoffs for Player I are shown first, followed by Player II.

Table 1. The mixed-motive game 'Leader'. Payoff for each player is determined by the actions (C or D) chosen by both players.

		Player II	
		C	**D**
Player I	**C**	2, 2	3, 4
	D	4, 3	1, 1

The abstract game of Leader describes a range of traffic and pedestrian interactions (among others). For example, suppose there are two cars waiting to cross a narrow bridge. The drivers must decide whether to drive (*D*) or concede right of way to the other person (*C*). If both concede then both will be delayed giving the outcome (2, 2). If they both decide to drive then a collision will occur resulting in the worst outcome for both (1, 1). If one decides to drive and the other wait, the 'leader' will receive a payoff of 4, while the other – the 'follower' – will be able to drive after the leader, receiving a payoff of 3. These concepts of leader and follower, and other concepts such as aggression/power, 'being good' and 'avoiding conflict' can be modelled as motive profiles that can guide the selections of actions in this scenario without requiring predefined goals.

Before we do this, we will modify the abstract game to capture several other related scenarios. Table 2 shows a range for payoff values that represent various government campaigns on road safety as increasing reward for being cautious and conceding right of way. Table 3 shows a range of payoff matrices with harsher penalties for collisions (such as higher insurance premiums or fines).

Table 2 Leader game with higher incentives for cautious driving.

		Player II	
		C	**D**
Player I	**C**	2,2	(3-3.8), 4
	D	4, (3-3.8)	1,1

Table 3 Leader game with higher penalties for collisions.

		Player II	
		C	**D**
Player I	**C**	2, 2	3, 4
	D	4, 3	(1-0), (1-0)

Agents Guided by Incentives and Implicit Motives

This section describes four agents that use Algorithm 1 with different motive profiles defined using the equations for power, achievement and affiliation motivation proposed by Merrick and Shafi (2011):

$$T_{res} = M_{ach} + M_{aff} + M_{pow} \text{ where}$$

$$M_{ach} = \frac{S_{ach}}{1 + e^{\rho^{+}_{ach}\left(M^{+}_{ach} - (1 - I_s)\right)}} - \frac{S_{ach}}{1 + e^{\rho^{-}_{ach}\left(M^{-}_{ach} - (1 - I_s)\right)}}$$

$$M_{aff} = \frac{S_{aff}}{1 + e^{\rho^{+}_{aff}\left(I_s - M^{+}_{aff}\right)}} - \frac{S_{aff}}{1 + e^{\rho^{-}_{aff}\left(I_s - M^{-}_{aff}\right)}}$$

$$M_{pow} = \frac{S_{pow}}{1 + e^{\rho^{+}_{pow}\left(M^{+}_{pow} - I_s\right)}} - \frac{S_{pow}}{1 + e^{\rho^{-}_{pow}\left(M^{-}_{pow} - I_s\right)}}$$

T_{res} is the resultant motivational tendency; I_s is the incentive value for a particular outcome. In this paper we use normalised payoff directly as incentive. This simplifies the definitions of task difficulty and risk associated with power, achievement and affiliation based on the assumptions that (1) higher payoff is generally associated with higher risk in power and affiliation motivation and (2) higher payoff is generally associated with higher difficulty in achievement motivation.

$M^{+}_{pow,ach,aff}$ are the turning points of approach components for each motive; $M^{-}_{pow,ach,aff}$ are the turning points of avoidance for each motive; $\rho^{+}_{pow,ach,aff}$ are gradients of approach; $\rho^{+}_{pow,ach,aff}$ are the gradients of avoidance; $S_{pow,ach,aff}$ are the relative motivation strengths of power, achievement and affiliation within the individual agent. The different motive profiles are created by setting and randomising the motive parameter values within different ranges as shown in Table 4. Agents use these parameters to compute the motivation value of each possible game outcome, then select the action (*C* or *D*) corresponding to the most highly motivating outcome.

Table 4. Motivation ranges for different motive profiles.

	Motive Profiles			
Param	**Leader-ship**	**Power**	**Achieve ment**	**Affili-ation**
S_{ach}	2	1	2	1
M^{+}_{ach}	-0.1	0.2	[0.2, 0.4]	0.2
M^{-}_{ach}	0.2,	0.5	[0.5, 0.7]	0.5
S_{aff}	1	1	1	2
M^{+}_{aff}	0.3	0.4	0.3	[0.3, 0.5]
M^{-}_{aff}	0.1	0.1	0.1	[0, 0.2]
S_{pow}	2	2	1	1
M^{+}_{pow}	[0.6, 0.8]	[0.8, 1.0]	0.9	0.9
M^{-}_{pow}	[0.95, 1.2]	[1.1, 1.3]	1.2	1.2

Results

Simulations of the agent with the leadership profile playing one hundred games from Table 1 against each of the three other agents are shown in Figure 5. The results show that agents with different motive profiles behave differently, indicating that different motive profiles affect the decisions they make, implying that the agents have different goals even though no explicit goals have been represented in the agents. Figure 6 shows that the agent with the leadership profile generally elected to drive (*D*). In contrast, the achievement motivated agent tended to wait, leading to a successful lead-follow interactions in almost all cases as shown in the first bar. In contrast, the more aggressive power motivated agent tended to drive more often. This resulted in a high number of collisions as shown in the second bar. The affiliation motivated agent also tended to drive in all instances, but its motivation was different. The affiliation motivated agent is motivated to seek low reward and avoid conflict with those who want higher reward. Ironically in this kind of situation, affiliation motivation results in poor performance (collisions). In motivational psychology, this sort of under-performance is considered characteristic of individuals with high affiliation motivation and lower achievement and power motivation (Heckhausen and Heckhausen, 2008).

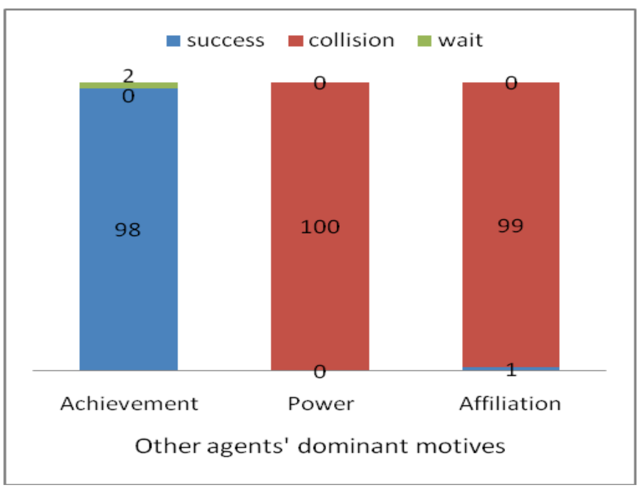

Figure 5. Outcomes of an agent with a leadership profile playing 100 games against each of three agents with the achievement, power and affiliation profiles respectively.

Simulations of the agent with the leadership profile playing Table 2 games against each of the three other agents are shown in Figure 6. Three Table 2 games with different payoff values are played one hundred times against each agent. The results show that when the payoff for waiting increases (caused by a government campaign on road safety for example) the agents adapt their behavior. In addition, different agents adapt differently. The agent with the leadership profile tends to wait more often as the payoff to wait increases. In contrast, the power motivated agent starts to wait on occasion, but still tends to drive most often. Therefore, we can see that with the leader versus power agent (Figure 7(a)), the successful outcomes are proportional to the payoff. The demonstration shows that our model of implicit motives permits agents to change their implied concrete goals as the environment changes.

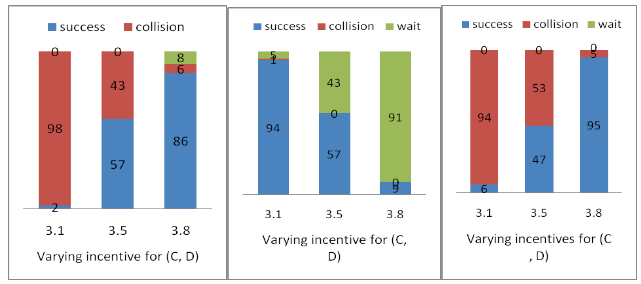

Figure 6. Outcomes of an agent with a leadership motive profile versus agents with strong (a) power (b) achievement (c) affiliation motive profiles, in Leader games with varying payoffs.

Where the power and leadership motivated agents were most affected by the 'government advertising campaign' affiliation motivated were most noticeably affected by the change in penalty for collisions as shown in Figure 7. Affiliation motivated agents tend to wait more often as the incentive for both agents driving decreased. Due to this, the number of collisions decreases as the punishment is increased. However, even with the highest penalties applied, there are still more collisions overall than with other types of agents. This experiment also demonstrates that agents with different implicit motives are able to respond to different types of changes, because they have a preference for different kinds of incentives. In addition, they did this in the absence of information about concrete goals.

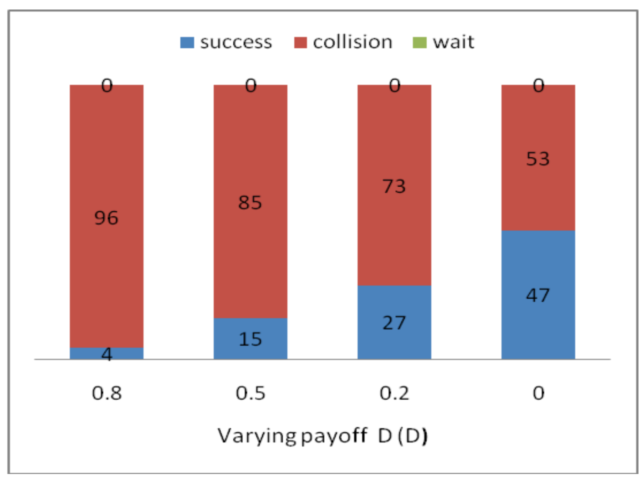

Figure 7. Outcomes of leader versus affiliation agent with increased punishment for collisions.

Summary

If we consider the Leader game in light of both the literature and the experimental results, then we see implicit goals, both abstract and concrete: abstract goals such as "being good", lead to concrete goals such as "avoiding collisions"; and abstract goals such as "being a leader" lead to concrete goals such as "driving first". The algorithms show that we do not need to represent either the concrete goals or abstract goals explicitly to achieve this behaviour. Rather, by representing incentives and different implicit motives we can model agents with different 'implied' abstract goals that influence their behaviour in different ways.

We summarise the literature and our own results in the context of Figures 2 and 3 with a spectrum of approaches that provide stronger to weaker representations of goals. These approaches complement goal-driven behavior or permit action in the absence of goals.

1. Strong models use domain-specific reward functions in learning and represent concrete goals in problem solving.
2. Weaker models reason in the absence of domain specific goals or rewards by calculating abstract goals or rewards based on the state of the external environment and the agent's experiences in its environment. These systems may or may not record goal structures in their internal representation. More importantly, in these systems, goal-like structures are flexible and intermediate states rather than provide the starting point for reasoning.
3. Finally, at the end of the spectrum, models of implicit motives can be thought of as reasoning in the absence of goals because they model only a preference for certain kinds of abstract implicit goals, rewards or incentives and not a goal itself. These models can still lead to action-selection and a type of emergent goal-directed behavior without explicit goal structures in the representation.

The distinction between (1) and (2) is similar to the Foner and Maes (1994) distinction between goal-driven and world-driven attention focus. Category (1) covers models that explicitly represent concrete goals, while category (2) covers models that explicitly represent abstract goals. In contrast, category (3) refers to models that define only a preference for certain kinds of abstract goals. In both categories (2) and (3) concrete goal-directed behavior is an emergent property. This emergent goal-driven behavior is critical for cognitive systems that can use 'idle time' effectively by continuing to monitor their environment to discover and pursue self-generated goals that extend or improve their knowledge base or skill set.

References

Baranes, A. and Oudeyer, P.-Y., 2010. Maturationally-constrained competence-based intrinsically motivated learning, IEEE International Conference on Developmenta and Learning, Ann Arbor, Michigan

Berlyne, D., 1960. Conflict, arousal and curiosity. McGraw-Hill, New York.

Braubach, L., Pokahr, A., Moldg, D., Lamersdorf, W., 2005. Goal representation for BDI agents, Second International Workshop on Programming Multiagent Systems: Languages and Tools, p 9-20.

Dignum, F. and Conte, R., 1998. Intentional agents and goal formation, Intelligent Agents IV: Agent Theories, Architectures and Languages. Springer, pp. 231-243

Fonar L and Maes, P. 1994. Paying attention to what's important: using focus of attention to improve unsupervised learning, Third International Conf on the Simulation of Adaptive Behaviour

Heckhausen, J. and Heckhausen, H., 2008. Motivation and action. Cambridge University Press, New York.

Horvitz, E., Apacible, J., Sarin, R. and Liao, L. 2005. Prediction, Expectation, and Surprise: Methods, Designs, and Study of a Deployed Traffic Forecasting Service, Proceedings of the Conference on Uncertainty and Artificial Intelligence, AUAI Press.

Itti, L. and Baldi, P. 2004. A Surprising Theory of Attention, IEEE Workshop on Applied Imagery and Pattern Recognition.

Kaplan, F. and Oudeyer, P.-Y., 2003. Motivational principles for visual know-how development. In: C.G. Prince et al. (Editors), Proceedings of the 3rd international workshop on Epigenetic Robotics : Modelling cognitive development in robotic systems, Lund University Cognitive Studies, pp. 73-80.

Kohonen, T., 1993 Self-organisation and associative memory, Springer, Berlin.

Langley, P., Laird, J.E., Rogers, S. 2008. Cognitive Architectures: Research Issues and Challenges, Cognitive Systems Research.

Maher, M.L., Poon, J. and Boulanger, S. 1996. Formalising Design Exploration as Co-Evolution: A Combined Gene Approach, in J.S.Gero and F. Sudweeks (eds) Advances in Formal Design Methods for CAD, Chapman & Hall, pp 1-28.

Marinier, R. and Laird, J. E. (2008). Emotion-Driven Reinforcement Learning. CogSci 2008, Washington, D.C

Marsland, S., Nehmzow, U., Shapiro, J, 2000 A real-time novelty detector for a mobile robot, In EUREL European Advanced Robotics Systems Masterclass and Conference.

McClelland, D. C., 1975. Power: the inner experience. Irvington, New York

Merrick, K., and Maher, M. L 2009. Motivated reinforcement learning: curious characters for multiuser games, Berlin, Springer

Merrick, K. and Shafi, K., 2011. Achievement, affiliation and power: motive profiles for artificial agents. Adaptive Behavior, 9(1): 40-62.

Ranasinghe N. and Shen,W-M, 2008. Surprise-Based Learning for Developmental Robotics. In Proc. 2008 ECSIS Symposium on Learning and Adaptive Behaviors for Robotic Systems, Edinburgh, Scotland.

Russell, S. J. and Norvig, P. 1995. Artificial intelligence: a modern approach, Prentice Hall, Englewood Cliffs, NJ

Saunders, R., 2001. Curious design agents and artificial creativity. PhD Thesis, University of Sydney, Sydney

Saunders, R. and Gero, J.S., 2004. Curious agents and situated design evaluations. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 18(2): 153-161

Schmidhuber, J. 1991 A possibility for implementing curiosity and boredom in model-building neural controllers. In The International Conference on Simulation of Adaptive Behaviour: From Animals to Animats pp 222-227

Singh, S., Barto, A.G. and Chentanez, N., 2005. Intrinsically motivated reinforcement learning. In: L. Saul, Y. Weiss and L. Bottou (Editors), Advances in Neural Information Processing Systems (NIPS). The MIT Press, Vancouver, pp. 1281-1288

Stanley, J. C., 1976. Computer simulation of a model of habituation. Nature 261:146-148