Noname manuscript No.
(will be inserted by the editor)

Biasing Behavioral Activation with Intent for an Entertainment Robot

the date of receipt and acceptance should be inserted later

Abstract Deliberate control of an entertainment robot presents a special problem in balancing the requirement for intentional behavior with the existing mechanisms for autonomous action selection. We propose that the intentional biasing of activation in lower-level reactive behaviors is the proper mechanism for realizing such deliberative action. In addition, we suggest that directed intentional bias can result in goal-oriented behavior without subsuming the underlying action selection used to generate natural behavior. This objective is realized through a structure called the intentional bus. The intentional bus serves as the interface between deliberative and reactive control by realizing high-level goals through the modulation of intentional signals sent to the reactive layer. A deliberative architecture that uses the intentional bus to realize planned behavior is described. In addition, it is shown how the intentional bus framework can be expanded to support the serialization of planned behavior by shifting from direct intentional influence for plan execution to attentional triggering of a learned action sequence. Finally, an implementation of this architecture, developed and tested on Sony's humanoid robot QRIO, is described.

Keywords behavior-based robotics · deliberative control · reactive-deliberative control

Address(es) of author(s) should be given

1 Introduction

As robots' interactions with humans become increasingly prevalent both in a service capacity [7] and in personal interaction roles [6], it becomes important to address the problem of incorporating intentional action into these systems in a natural way. This is an exceptionally challenging proposition for entertainment robots because the intentional control needs to provide for goal directed action without subsuming the existing action selection mechanisms used to provide the robot with "personality" [21]. In the case of a humanoid, such as Sony's QRIO shown in figure 1, the problem is further compounded by the expectations placed upon anthropomorphic robots by the user. In order for a humanoid's actions to be perceived as natural, its deliberative capabilities need to extend beyond that of simple plan execution and move closer to the range of deliberative capabilities afforded to humans.

Intentional behavior, in the context of this work, can be defined as any behavior that is deliberately invoked in order to achieve a high-level goal. High-level goals, in this regard, are goals that are not explicitly encapsulated within the reactive behaviors used by the system. High-level goals may be realized in a number of ways including the activation of a particular sequence of behaviors or the inhibition of certain behavior. Examples of high-level goals that may exist in an entertainment robot can range from the simple (e.g. moving to a particular room at a particular time) to the increasingly complex (e.g. searching for objects in the environment, collecting them, and delivering the objects to the user).

As stated earlier, the primary goal of this work is to incorporate these types of intentional activities in a natural manner. Since natural behavior can be subjective, the following will serve as the working definition for this article: Natural intentional behavior is deliberative behavior that resembles the type of intentional capabilities that are exhibited by humans. The types of intentional control exhibited by humans have been divided into several categories by psychologists [22]:

1. Tasks involving planning or decision making
2. Tasks involving troubleshooting
3. Tasks containing novel sequences of actions
4. Tasks that are dangerous or difficult
5. Tasks that require the overriding of habitual responses

While intelligent systems may be able to benefit from handling all the deliberative tasks described above, the three most relevant, at least from the viewpoint of an entertainment robot, are tasks requiring planning, tasks requiring novel sequences of actions, and tasks requiring the overriding of habitual responses. As it is unlikely that an entertainment robot will need to perform dangerous tasks or be required to troubleshoot, these capabilities will not be addressed in this article.

While many examples of systems in which deliberative control in the form of plan generation and execution exists (e.g., [20][23]), these systems rarely allow for deliberative control to be expressed beyond sequences of actions generated by the planner. For an entertainment robot, however, the range of

Fig. 1 Sony's entertainment robot QRIO

control provided by a deliberative system should allow for more then merely sequenced plan execution. Deliberative control should also serve as a source of novel actions and novel plan execution. It addition, deliberative control should not only serve as a means of expressing action, but also as a means of suppressing action in situations where they may be inappropriate.

Finally, in any system that wishes to express intention in a natural manner, a plan must be able to be expressed differently based on the importance of its goals. An example of this can be seen when a person has a goal of attending a meeting in a room down the hall. If time before the meeting is sufficient, one may stop to chat with colleagues while on their way to the meeting room. If late to the meeting, such interaction may be skipped in an effort to reach the meeting quickly. In the case where there was sufficient time to make the meeting, the high-level goal of attending the meeting was of low priority. In the other case, the same goal was of higher priority and thus results in different behavior. Similar, fine-grained expression of intentional control in an entertainment robot is a crucial capability that must be addressed if the robot is to go beyond mere plan execution.

This work proposes that intentional bias of the action selection mechanism used by entertainment robots such as QRIO is the proper means by which to express such deliberate control in a natural manner. Because of the delicate balance required between a robot's behavior serving in an entertainment and service capacity, it is inappropriate for these functions to exist in isolation. One does not want a robot meant to serve as a companion to lose all semblance of personality or related entertainment characteristics when it is pursuing high-level goals. Similarly, it is not desirable for an entertainment robot to exist solely in the context of satisfying its immediate needs in a purely reactive manner.

Our research presents a mechanism for providing uniform intentional control of the types described above. This mechanism, called the intentional bus, serves as a gateway between the reactive and deliberative layers in the architecture. The intentional bus converts high-level goal-oriented tasks generated by the deliberative system into intentional signals that serve to bias behavioral activation in the reactive layer. These biases serve as a means of directing the robot towards accomplishing the goals of the deliberative system without overriding the existing action level computations used to generate natural behavior. In addition, we show that this interface is general enough to provide interaction beyond that of intentional bias by expanding it to handle attentional triggering of learned sequences of behavior. Finally, we present an implementation and results for this form of deliberative control on Sony's humanoid QRIO.

The remainder of the paper is composed as follows: Section II provides an overview of related work on hybrid reactive-deliberative architectures, psychological evidence for intentional bias, and an overview of the action-selection architecture for the entertainment robot QRIO used in this work. Section III describes the intentional bus developed for this research and its interactions with the deliberative and reactive architectural layers. Section IV extends the intentional bus design to allow for attentional triggering of learned serial actions. Section V describes the deliberative architecture developed using intentional bias to provide high-level control for an entertainment robot. Section VI provides an evaluation of the use of intentional bias for control of QRIO. Finally, section VII concludes the paper and presents an agenda for future work.

2 Related Work

2.1 Hybrid Reactive-Deliberative Architectures

Hybrid reactive-deliberative systems have been successfully deployed for many years, serving as a mechanism for combatting the shortfalls of purely reactive and purely deliberative architectures. Arkin and Gat provide excellent overviews of a number of hybrid architectures in [2] and [15]. In these architectures, the deliberative layer is usually coupled to the reactive layer in one of three manners: as a mechanism for generating actions at a differing time scale than the reactive layer; as a means to guide reaction; and through the direct coupling of reaction and deliberation, each guiding the other [18]. In most cases, the role of the deliberative system is to plan which actions should execute at a particular time to achieve a goal. In many cases, the interaction between the deliberative layer and the reactive layer is mediated by a sequencing layer which selects the portions of the plan that should execute at a given time. The sequencer then activates the behaviors required to realize the next step in the plan.

In most architectures, this activation is almost always binary, without support for an increase or decrease in the likelihood of behavioral activation. Some work has investigated alternative means of combining deliberative information directly into the reactive layer [24][27]. These plans provided a navigational bias into a reactive controller, which served as a means of influencing the expression of the currently executing behavioral assemblage while still allowing other reactive behaviors, such as obstacle avoidance, to be active. The plans used in their work were not incorporated into the behavioral selection process itself, however.

Numerous other means of incorporating deliberative information into the reactive layer have been investigated. In almost all cases, this influence is expressed as the binary selection of reactive behavior and do not allow for variable levels of activation as proposed in this article. An example of such a coupling include Rosenblatt's DAMN architecture [25] which provides a mechanism where behaviors in the traditional sense, as well as planners instantiated as behaviors, cast votes as to the appropriate action the robot should take. The arbiter in this architecture combines these votes to select the appropriate action to take. Significant challenges exist in specifying voting semantics making arbitration between different behaviors and fine-grained deliberative control difficult, however.

Carpenter *et. al.* developed an architecture for robot soccer in which deliberate advice from a coach is integrated via the addition of suggested behaviors to the set of executable behaviors [8]. The effect of this advice, however, is highly dependent on the scoring function used by the behavioral arbiter and appears unable to be influence behavioral selection in a non-binary manner.

2.2 A Psychological Basis for Deliberative Control of Reactive Behaviors

One particularly compelling model for the deliberative control of automatic actions in humans, developed by Norman and Shallice, is the supervisory attentional system (SAS) [22]. Created to account for lapses of action in routine activity, they propose the SAS acts as the controller over the expression of automatic behaviors within humans. An overview of the interaction of the SAS with the underlying behavioral layer can be seen in figure 2. Their overall behavioral model can be divided into two major interacting components: the contention scheduling mechanism and the supervisory attentional system.

The contention scheduler involves multiple sets of simple, well-defined actions called schemas. Schemas can be atomic or an ordered series of schemas and are activated by certain triggering conditions perceived in the environment. While it is beyond the scope of this paper to discuss contention scheduling in detail, there has been significant work in formalizing models of contention scheduling [10] [11] as well as evaluating its feasibility as an action-selection mechanism for robots [1] [16].

The contention scheduling system serves as a means of providing reactive control for routine actions. In the case of novel sequences of actions that are not known *a priori*, the contention scheduling system alone is insufficient to express such behavior. Norman and Shallice [22] propose that

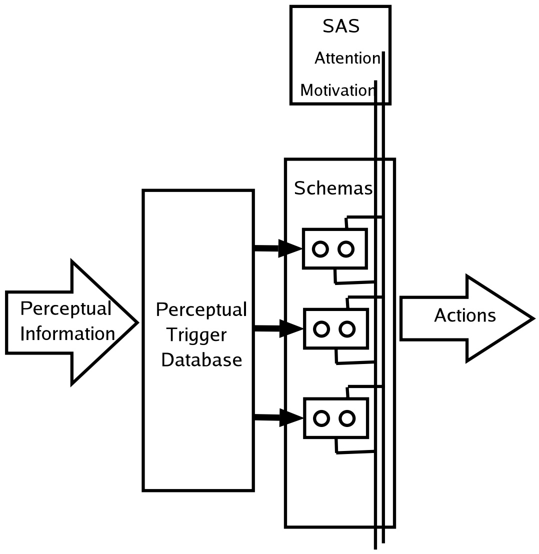

Fig. 2 Norman and Shallice's model of the role of the SAS (after [22])

the SAS is the component that provides the additional source of control needed to exhibit novel actions. They suggest that the SAS interacts with the reactive contention scheduling layer through the use of motivational and attentional threads. These threads bias the activation of schemas in the contention scheduler without activating them directly. These biases can allow for the activation of novel sequences of behavior even if that sequence did not already exist in the contention scheduler. How the SAS actually does this is to date still ill-defined, although Shallice and Burgess have examined possible mechanisms by which the SAS may generate and evaluate plans [26]. They do not, however, talk about the influence of the attentional and motivational threads with the lower-level reactive layer in much detail.

While it has been proposed [2] that Norman and Shallice's model of willed behavior may provide a suitable guide for integration of reactive and deliberative control in robotic systems, little has been done to investigate such mechanisms. Garforth *et. al.* appear to have been alone in doing so thus far. They looked at using neural mechanisms to mimic processes of the supervisory attentional system for a simulated robot [13] [14], investigating the suppression of a particular behavior in a simple foraging task using a previously trained neural network. The controller in this architecture appears to be limited to the suppression of a particular behavior and cannot generate novel behavioral sequences itself. In addition, their architecture does not appear to support non-binary suppression.

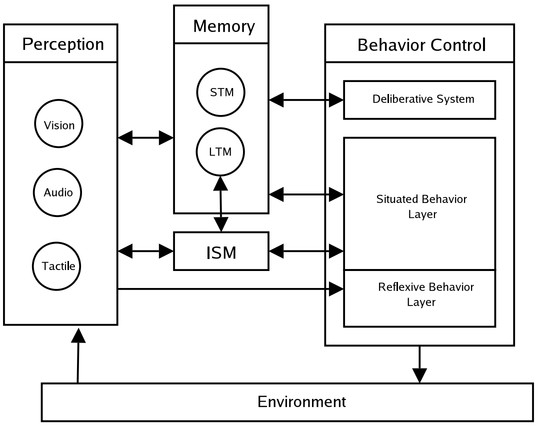

Fig. 3 Overview of the EGO architecture. Short-term memory and Long-term memory are denoted by STM and LTM respectively.

2.3 Action Selection and Behavioral Representation in the EGO Architecture

Action selection, the problem of determining which action to perform among many possible conflicting actions, has been studied extensively in both the agents and robotics communities. Numerous models have been produced including ethologically-guided [4] and spreading activation models[19]. These efforts, however, do not examine the role of intention in action selection.

For this work, we use the action selection mechanism incorporated within the EGO architecture [12]. The EGO architecture, or Emotionally GrOunded architecture, utilizes the OPEN-R framework originally designed for the quadruped robot AIBO and later expanded to the humanoid QRIO, the research platform for this work. An overview of the components within the EGO architecture is shown in figure 3. In the EGO architecture, behaviors are organized in a tree structure according to their conceptual relationship with each other [3]. For example, a high-level soccer behavior may be decomposed into several low-level, reactive behaviors such as finding and kicking a ball 4.

In the EGO architecture, behaviors maintain an activation level and set of resources associated with them. Every cycle, most behaviors calculate their activation levels after which a selection process occurs. In the action selection process, the behavior with the greatest activation level for which there are available resources is selected for execution . If there are remaining hardware resources, this process is repeated until there are no more free resources. This allows for the concurrent execution of non-conflicting behaviors to occur.

In the EGO architecture, the activation level calculation is a result of an ethologically-based homeostatic action selection mechanism [3]. The homeostatic control module in QRIO attempts to keep a number of internal state variables, such as desires for interaction, rest, and activity, within certain bounds. By evaluating the levels of these particular variables as well as external stimuli (releasing factors), actions can be selected by increasing the

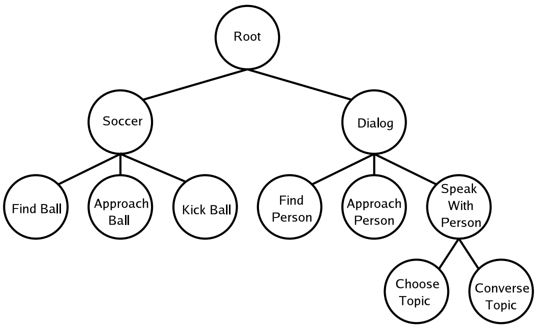

Fig. 4 Example behavioral tree in the EGO architecture containing two main behaviors: Soccer and Dialog

activation levels of behaviors that will satisfy the robot's internal drives. The further a particular internal state variables is out of the specified range, the greater its influence on action selection.

More precisely, the activation level for the behaviors in the EGO architecture are calculated as:

$$B_v = \beta M_v + (1 - \beta) R_v, \tag{1}$$

where B_v is the activation level, M_v is the motivational value based on the robots internal state, and R_v is the releasing value. The releasing value is be specified as:

$$R_v = \alpha \Delta S + (1 - \alpha)(S + \Delta S), \tag{2}$$

where S is the current satisfaction as measured by the internal state variables and ΔS is the expected change is satisfaction if the particular behavior were to become active. A detailed discussion on the activation level computation and action selection within the EGO architecture can be found in [17].

To reiterate, our hypothesis is that through the intentional bias of activation levels such as described here, goal-oriented behavior can be added to the system without compromising the strength of the existing action selection mechanism.

3 The Intentional Bus as a Mechanism for Deliberative Behavioral Biasing

In order to combine the execution of reactive behavior with goal-oriented behavior in a natural manner, a mechanism for providing, monitoring, and maintaining intentional control is necessary. The intentional bus is the component that provides these services and allows the deliberative system to interact with the reactive behaviors in a coherent manner. An overview of the interaction between the deliberative layer and the lower level behavioral layer appears in Figure 5. The intentional bus provides three major functions:

1. Monitoring and reporting the status of the underlying behaviors.
2. Biasing the activation levels in the reactive layer via intentional threads.

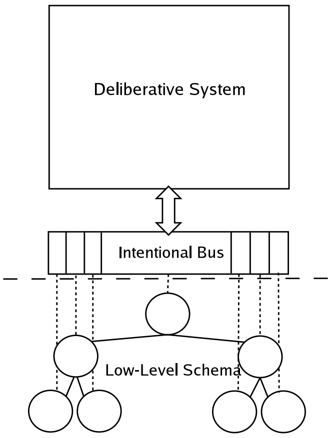

Fig. 5 Overview of the interaction between the deliberative system, intentional bus, and lower level reactive behavior

Schema Name	Dance	Sing	......	Approach	Greet
State	Running	Stopped	......	Stopped	Stopped
Completion	False	False	......	True	False
Activation	100	50	......	10	10
Intention	60	0	......	0	0
Attentional Flag	False	False	......	False	False
Process ID	1	2		22	23

Fig. 6 The state information contained within the intentional bus with sample values

3. Maintenance of the intentional bias in response to changing activation levels.

3.1 Monitoring the Reactive Layer

The intentional bus serves as a repository of information about the underlying reactive level for use by a deliberative system. In addition, this state information is used by the intentional bus itself for calculating the appropriate intentional bias to send to behaviors. The bus stores information pertaining to the state of all behaviors (running, stopped, etc.), activation levels of all behaviors, and the intentional levels of all behaviors. This information is derived from two different locations: Intentional data comes from the deliberative layer while all other information is received from the reactive layer. An example of the state information stored and utilized by the intentional bus is depicted in figure 6. In addition, figure 7 shows the origins and destinations for all data in the bus as well as the computations performed on that data by the bus.

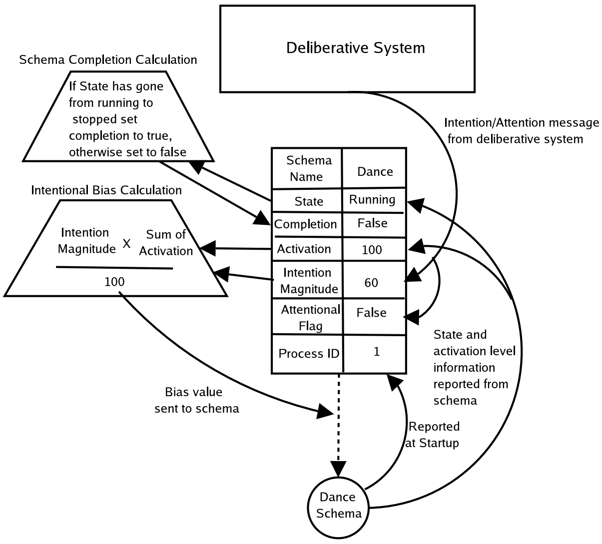

Fig. 7 Dataflow in the intentional bus. Arrows show origin and destination of the data. Trapezoids show computations internal to the intentional bus.

Algorithm 1 Biasing Behavioral Activation for Behavior s at time t

$activation_{total} = 0$
for each behavior s **do**
    $activation_{total} = activation_{total} + activation_{s,t-1} - intention_{s,t-1}$
end for
$intention_{s,t} = \frac{activation_{total} * bias}{100}$
$activation_{s,t} = (activation_{s,t-1} - intention_{s,t-1}) + intention_{s,t}$
record intention type (bias or attentional trigger)

3.2 Biasing Activation Levels via Intention

The second function of the intentional bus, the biasing of behavioral activation levels, is initiated upon request from the deliberative system. This request is composed of three parts:

1. s: the behavior to bias
2. m: the magnitude of the bias
3. a: an attentional flag, set to indicate if the bias should be an attentional trigger (described in detail in the next section)

The magnitude of the bias is calculated by treating m as a percent of the total activation of all behaviors *minus* the total intentional bias of the system, such that,

$$b = \frac{m \sum (a_i - b_i)}{100} \tag{3}$$

where a_i and b_i represents the activation level and bias for behavior i. Other means of calculating the bias can be substituted without loss of generality, however. The resulting bias value is placed on the bus where it is sent to

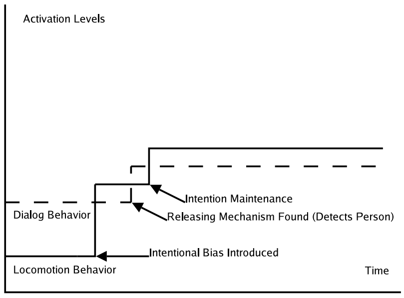

Fig. 8 An idealized graph of activation values for two behaviors showing how intention can bias the activation level of a behavior and how the intentional bus modulates the intention in response to outside influences. In this example, the locomotion behavior is biased via intention enough to cause it to run. While the locomotion behavior is executing, however, the releasing mechanism for dialog behavior is found which causes its activation level to increase. The intentional bus observes this and adjusts locomotion behavior's bias appropriately.

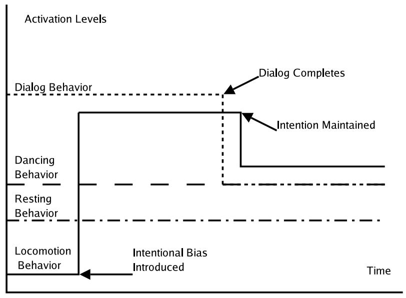

Fig. 9 An idealized graph of activation values for four behaviors demonstrating how intentional bias does not necessarily override existing behavioral expression. In this case the bias is not great enough to make the locomotion behavior run. The bias remains, however, and when the robot finishes conversing with a user, the bias to the locomotion behavior now allows it to run.

the appropriate behavior, s, via the intentional threads and results in the activation level increasing or decreasing. The bias on the activation level of the behavior can have a large number of effects ranging from certain execution for values of $m > 100$, slight influence towards activation when $0 \leq m \leq 100$, to inhibition of the behavior when $m \leq 0$. An unbounded increase in activation levels is avoided by subtracting the activation contributed by intention. Ensuring that activation levels for individual behaviors remain bounded is the responsibility of the underlying action selection mechanism. Algorithm 1 depicts an overview of the process of behavioral bias.

The values of m can be derived from a number of sources. These sources may include time constraints for plan completion where goals that must be accomplished quickly are assigned higher values of m. Values of m may originate from the preferences for certain activities either assigned to or learned by the robot. In these cases the robot may assign lower values of m when performing actions that it does not enjoy. Another potential source of bias magnitude could simply be the user who can inform the robot of the importance of a particular goal. For this work, however, the origin of the bias magnitude is not as crucial as the effect of the bias on the system itself.

After intentional biasing has occurred, the bus begins monitoring the status of the biased behavior to determine when to stop the bias. When it has detected the behavior transition from stopped to running or running to stopped, intentional bias is adjusted as the situation calls for it. This monitoring process is detailed in algorithm 2.

Algorithm 2 State Monitoring

```
for each behavior s do
    if the status of s changed then
        if s stopped running then
            intention_s = 0
        end if
        if s started running then
            if s is being biased and bias type is attentional then
                intention_s = 0
            end if
        end if
    end if
    record behavior data (activation levels, state, etc.)
end for
```

3.3 Intention Maintenance

The final function of the intentional bus is the active maintenance of the intentional signals. As activation levels change in the system due to the termination and activation of behavior as well as the detection of releasing mechanisms in the environment, the intentional bias currently applied to a behavior may no longer remain at the appropriate level. In the event there is a change in the activation levels of the underlying behavior, the intentional bus uses the stored values of m and recalculates the intentional bias for each behavior and adjusts the bias accordingly. The maintenance process can be achieved executing algorithm 1 over each behavior each time a change in behavioral activation levels is detected.

Figure 8 shows a simple example in which the robot maintains intentional bias under fluctuating activation levels. In this example, the intentional bus is realizing a simple plan to walk to some location (using the locomotion behavior). While executing this plan the robot encounters a person nearby. This results in the activation level of the dialog behavior to increase (detecting the presence of a person is the dialog behavior's releasing mechanism). In this case, the intentional bus detects this change and actively maintains the intentional bias. As the bias magnitude is high, the robot continues walking toward its goal instead of speaking with the person present.

A second example of intentional bias that does not override the existing action selection can be seen in Figure 9. In this example, the intentional bus once again provides the bias necessary to execute a simple plan to walk to a location. In this case, however, the magnitude of the bias is lower then that used in the previous example. As a result, the robot continues to speak to the user despite the executing plan. When the robot is finished talking, the dialog's activation level falls and the bias to the locomotion behavior is modulated. Now, however, the activation of the biased locomotion behavior is high enough to allow that behavior to execute.

4 The Intentional to Attentional Shift for Routine Activity

One important capability for a robot that is designed for long-term interaction is the ability to learn routine activities from repeated intentional execution of plans. Routine activities in the context of this work are plans generated by the deliberative layer that are repeated many times. When a previously planned activity becomes a routine activity, the explicit plan execution and monitoring that occurs in the deliberative system can migrate to the lower-level action selection mechanism, freeing the resources that formerly had been dedicated to generating and executing such plans. We propose that this transition can occur via the use of an attentional signal to activate a learned behavioral sequence. Instead of the deliberative system executing an activity through the intentional bias of the underlying behaviors, an attentional signal is instead sent to the first behavior in the sequence in order to start a chain-activation of the requisite behaviors. This notion of initiating sequences of actions at a low-level via a high-level process has been touched upon by Bonasso and Kortenkamp [5] but never investigated fully.

Describing the specific learning mechanism used within the reactive layer for learning serial actions is beyond the scope of this paper, but details can be found in [9]. Other alternative designs are also possible. For the purposes of this article, it is assumed that such a mechanism exists and that the reactive layer can relay a message to the deliberative system indicating that a task has become routine. Upon receipt of this message, the deliberative system shifts from providing a continuous intentional influence to instead providing a short attentional signal to start the sequence of actions corresponding to that plan. This occurs through the intentional bus using the same mechanism that handles intentional signals. Once this signal from the low-level, reactive behavior has been received, the current sequence of actions can be marked as routine at the deliberative level. If the series of behaviors has in fact been learned, in the future the deliberative system can send a request to the intentional bus that an attentional message be sent to the first behavior in the sequence whenever that sequence is required to be re-activated. This attentional trigger differs from the intentional bias only in duration.

For an attentional trigger, the intentional bus only sends the bias until the behavior begins activation. After that, the behavior is triggered and the bias is set back to zero instead of being actively maintained at a specific level. The first several times the attentional signal is sent, however, it is desirable for the deliberative system to monitor the execution of the sequence to ensure the sequence has indeed been learned properly. In the case that it has, the deliberative system no longer interacts with the subsequent behavior in the sequence but instead allows the sequence to execute automatically at the reactive level.

An illustration of the intentional and attentional activity before and after a serial action is learned can be seen in figures 10 to 13. The serial behavior shown in the example is a simple soccer activity consisting of approaching and then kicking a ball. When executed by the deliberative system, each of the approach and kick behaviors are biased and maintained in turn (Figures 10 and 11). After this simple soccer plan has been executed by the deliberative

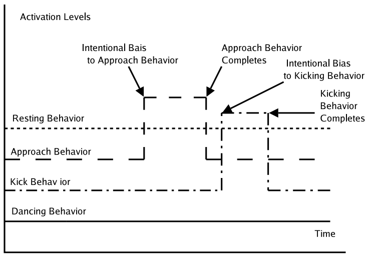

Fig. 10 An idealized graph of activation levels for a soccer behavior composed of two sequential behaviors before learned serialization.

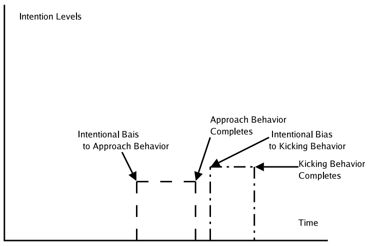

Fig. 11 An idealized graph of deliberately applied intentional values for a series of two behaviors before learned serialization.

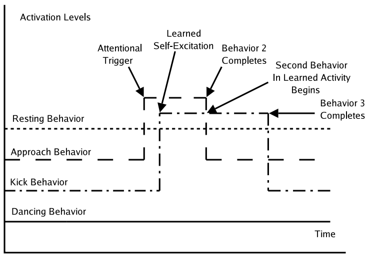

Fig. 12 An idealized graph of activation values for a learned routine soccer behavior. In this example, after the approach behavior becomes active, the kick behavior's activation increases via self-excitation in anticipation of being active next.

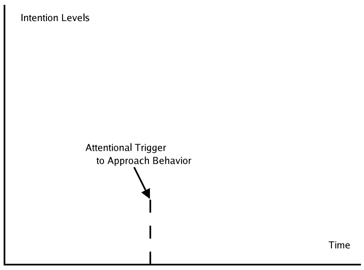

Fig. 13 An idealized graph of intentional activity showing an attentional trigger for a learned routine soccer behavior. The trigger causes approach behavior to become active and allows the complete learned routine to express itself automatically.

system many times it may become a routine activity encapsulated in the reactive behavioral system. Once this has happened, the intentional bus need only to provide an attentional trigger to initiate the learned approach, kick routine sequence (Figures 12 and 13).

5 Design and Implementation of an Architecture for Deliberative Control via Intentional Bias

An overview of the deliberative/reactive architecture utilizing intentional bias can be seen in figure 14. The deliberative system can be viewed as a collection of interacting subsystems. These subsystems provide mechanisms for planning, plan execution, state tracking, as well as the intentional bus for interfacing the deliberative layer with the reactive layer.

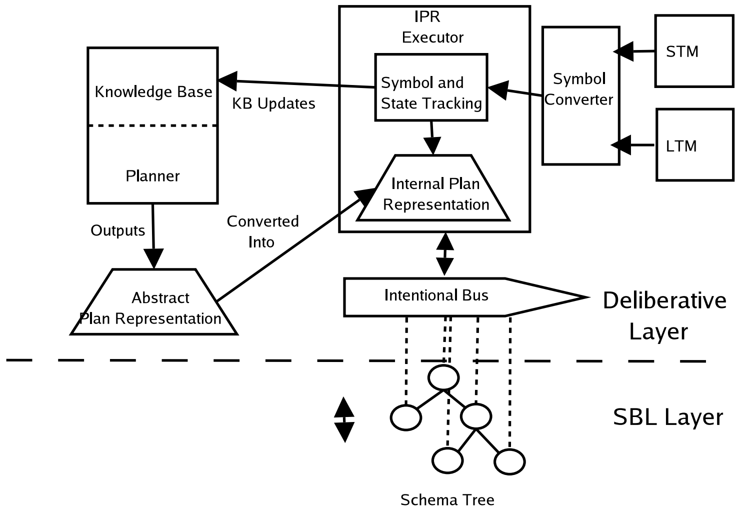

Fig. 14 Overview of the end-to-end deliberative/reactive architecture

5.1 Planner, Knowledge Base, and the Internal Plan Representation

Generation of plans is one of the primary functions of a deliberative system. To accomplish this task a suite of three components is utilized: the planner, the knowledge base, and a module to convert the output of the planner to a format usable by the robotic platform. The knowledge base contains facts and assertions about the world, the state of the robot, the tasks that must be accomplished, as well as knowledge concerning the different behaviors the robot can execute and their effect upon the world. The knowledge base also contains the information required for the generation of the bias values in a plan. This knowledge takes the form of scalar values representing user and robot preferences for tasks. Future work will look into means of acquiring these values over the lifetime of the robot using on-line learning algorithms. The planner uses the information contained in the knowledge base to generate plans based on the goals of the robot at a given time.

To generate plans for execution in this architecture, a hierarchical task network planner was used for the initial implementation, which outputs plans consisting of ordered sequences of actions that need to be accomplished in order to achieve the goals maintained in the knowledge base. This output can vary based on the specific type of planner being used. We do not use the planner's output directly so as not to tie the deliberative system to any one particular planner implementation but instead convert it into a generic format usable by the system. This format is called the Internal Plan Representation (IPR).

An IPR statement includes information about the world state for which a behavior should be biased, how much to bias the behavior, and any additional parameters necessary for the behavior's execution (e.g. the goal location or the object to interact with).

Algorithm 3 Plan Execution

```
for the next IPR statement i that is not executing do
  if state_i matches world state then
    if behavior_i is already running then
      resolve behavior priority conflicts
    end if
    if symbols in i are not assigned to external objects then
      bind objects in short-term memory to symbols in i
      for each IPR statement j do
        if symbols_i = symbols_j then
          bind objects assigned to symbols_i to symbols_j
        end if
      end for
    end if
    calculate intention from priority_i
    bias behavior_i
    set i to running
  end if
end for
```

5.2 Plan Execution

The IPR executor is the core module for the execution of plans generated by the deliberative system. Its functions include plan item selection, plan progress tracking, symbol binding, and behavior conflict resolution. The primary task for the IPR executor is the determination of applicable plan items based upon their state requirements and the current state of the world. This process occurs in several stages. The first phase selects the set of plan items which are not executing and whose state requirements currently match those of the world state. This set is then tested for potential behavior conflicts (described below). The plan items without conflicts are then bound to the proper symbols in memory. After binding they are then set to *running*, and a message is sent to the intentional bus indicating which behavior to bias and the associated bias magnitude as specified in the IPR statement. In addition, the message contains the appropriate flag designated by the plan item indicating if it is an intentional request requiring maintenance by the bus of the behavioral bias or if it is instead an intermittent attentional trigger used for a routine activity. The algorithm 3 depicts a high-level overview of the the process used for plan execution.

The second major task the IPR executor performs is the binding of symbols specified in the plan to their counterparts in the environment. This helps ensure that different plans and partial plans share the same target information despite belonging to different subtrees in the underlying behavioral tree.

Schema conflict resolution, the third function of the IPR executor, refers to the fact that multiple plan items may be in competition for a single behavior. For example, there may be two potentially active plan items which both require use of the robot's approach behavior but for two different objects. Only one of these plan items can be active at a particular time. The IPR executor uses a priority-based arbitration mechanism to resolve these conflicts where the priority for a particular plan item is represented by the magnitude of the intentional bias as specified in the IPR representation. Whenever a plan item is applicable, the IPR executor will first check that the appropriate behavior is not already in use by another plan item. If it is, it compares their bias magnitudes and selects the plan item with the highest magnitude. In the case of a tie, the plan item that currently has control of the behavior remains the executing behavior. Hardware resource conflicts (i.e. two competing behavior that each require use of the right arm) are handled by the underlying action-selection system. This allows multiple plans or partial plans to safely execute in parallel.

The final task of the IPR executor is tracking the plan's progress, notably, which plan items are running and which plan items have been completed. This is done by querying the intentional bus for the status of biased behaviors. When it has discovered a biased behavior has been completed, the bias is set to zero and the relevant state information in the knowledge-base is updated.

5.3 Intentional Bus

As described previously, the intentional bus serves as the interface between the deliberative system and the underlying reactive layer. The intentional bus stores a large amount of information about the underlying behaviors such as their state (ready, active, sleeping, etc.), their current activation levels, their process ids, and their current intentional levels. Messages sent from the reactive layer concerning behavior status and activation levels are routed to the intentional bus for determining behavior completion.

Upon receipt of a message from the IPR executor indicating a specific behavior to bias, the intentional bus stores the requested bias magnitude for later bias calculation as well as the flag indicating if it corresponds to an attentional trigger or intentional bias. In the case that the magnitude is zero, the intentional bias is immediately stopped for that behavior. At every timestep, the intentional bus examines the bias magnitude value for each behavior, and uses this value to calculate the appropriate intentional bias to send to the associated behavior as described in section 3. If the intentional bus has been requested to send an attentional trigger to a particular behavior, the next timestep the intentional level is set to zero.

6 Experiments

To demonstrate the ability of the implemented system to provide goal-oriented behavior compatible with the pre-existing action selection mechanisms, three experiments were conducted. The goal of these experiments was to verify that intentional bias can support the subset of human deliberative abilities suitable for producing natural behavior in an entertainment robot: traditional plan execution, behavioral suppression, and novel behavioral expression. One of these experiments was conducted as a numerical simulation, the other two experiments were conducted on the QRIO platform.

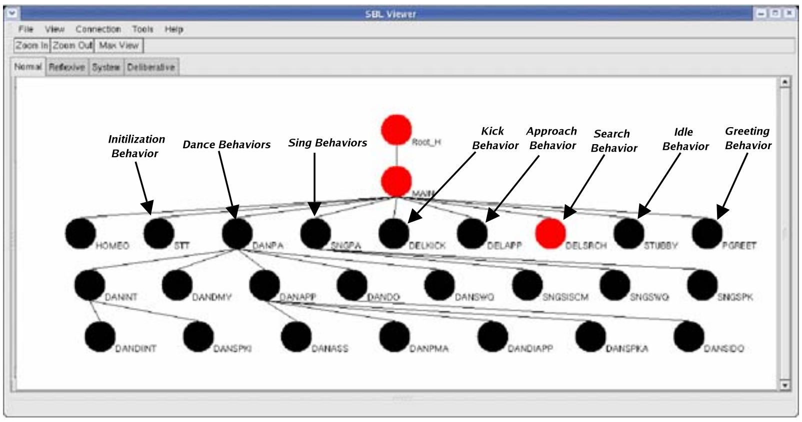

Fig. 15 Behavioral tree used in intentional experiments. The bold-italic labels indicate the behavior associated with that node and any of the node's children if applicable.

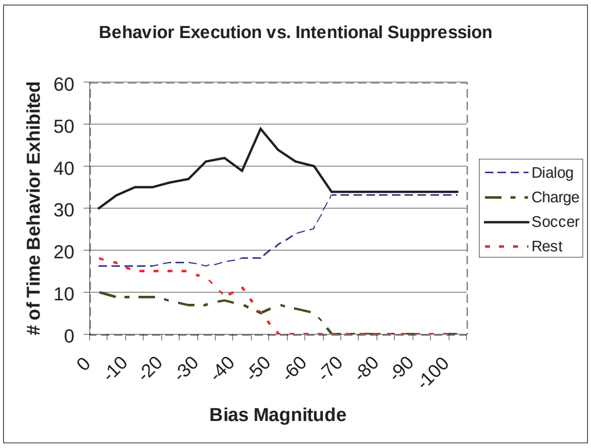

Fig. 16 Behavior exhibited as a function of the amount of intentional bias on the charge and rest behaviors. Demonstrates how intentional bias can provide variable amounts of behavioral suppression.

6.1 Behavioral Suppression via Intentional Bias

In the first experiment we examine the ability of the system to exhibit behavioral suppression through the use of intentional bias. The backstory for the scenario can be described as follows: the owner of an entertainment robot is having some friends over to show it off. When the guests arrive, she doesn't want the robot sitting around doing nothing but rather entertaining the guests. She communicates this request to the robot and as a result, the robot instantiates a plan in which its 'boring' behaviors are suppressed. The goal of this experiment is to examine how the suppression of behavior via negative intentional bias will affect the action selection of the robot.

In this experiment the robot has four high-level behaviors: a dialog behavior, a soccer behavior, a resting behavior, and a battery charging behavior. A numerical simulation was run to compute the activation levels of these behaviors based on the homeostatic action selection mechanism reviewed in section 2.3 and [17]. In the experiment, a variable amount of negative intentional bias (suppression) was applied to the unentertaining rest and charge behaviors. At each timestep in which a behavior was not activated, the behavior with the highest activation level was run. The rest and soccer behaviors ran for 5 timesteps each time they were activated, the charge and dialog behavior had a duration of 10 timesteps. The simulation was run for 500 timesteps total. Twenty-one trials were run where the intentional bias for the rest and charge behaviors ranged from 0 to -100 in increments of 5.

Figure 16 shows a graph of the expressed behavior in relation to the magnitude of behavioral suppression. The behavioral expression of the core homeostatic action selection mechanism can be seen when the intentional bias is equal to zero. In this case, the robot rests when it is tired, chats when it desires interaction, etc. As the amount of intentional bias decreases below zero, the frequency of the 'boring' behaviors decreases until they eventually disappear (at $m = -65$). Important to note, is the fact that intentional bias does not necessarily turn off behavior, but instead actively discourages its use. The result of this suppression, in the case of the charge behavior, is that the robot's battery level must get lower before the robot's desire to recharge itself becomes great enough for the charge behavior to become active. This demonstrates that intentional bias can provide fine-grained deliberative control via behavioral suppression without necessarily overriding the core action selection mechanism.

6.2 Plan Execution via Intentional Bias

The second experiment looks purely at the ability to rigidly follow plans by executing them through the addition of large amounts of intentional bias ($m > 100$). The goal of this experiment is to demonstrate that intentional bias of reactive behavior can ensure plans can get executed for very important tasks or in situations in which traditional plan execution is desired. This experiment was conducted on QRIO platform. For this experiment, a behavioral tree was used consisting of several behaviors including dancing, singing, approaching, searching, kicking, a reflex action to turn towards loud noise, and a greeting action (figure 15).

To demonstrate the ability of intentional bias to serve as a mechanism for executing plans, a simple plan was generated to play soccer. This plan consisted of the three steps: searching for, approaching, and then kicking a pink ball. While this plan is being executed by QRIO, the operator calls out to QRIO to try to get it to interact with him. The operator then shows QRIO a green ball, the releasing mechanism for the greet behavior. The greet behavior, if activated, causes QRIO to stop, wave, and then sit down for a few moments to interact with the user. It is proposed that through the modulation of the intentional bias in reaction to events such as this, the underlying action-selection can be overridden when necessary.

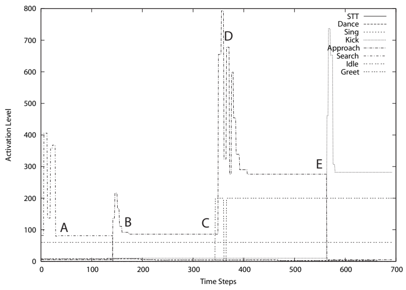

Fig. 17 Activation levels trace during the execution of a high intentional magnitude plan: A) Search behavior is active. B) Approach behavior becomes active. C) Releasing mechanism for greet behavior is found and its activation is increased. D) Intentional bus maintains appropriate intention levels keeping approach active. E) Kick behavior becomes active.

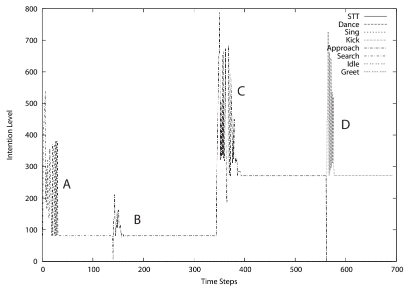

Fig. 18 Intention levels trace during the execution of a high intentional magnitude plan: A) Intentional bias is sent to search behavior. B) Intentional bias is stopped for search and started for approach behavior. C) Intentional bias is increased by intentional bus due to changes in behavior activation levels. D) Intentional bias to approach is stopped and bias to kick behavior begins.

(i) **(ii)** **(iii)**

Fig. 19 Depiction of high intention soccer plan. i) QRIO approaches the ball ii) The operator tries to distract QRIO with the green ball (the releasing mechanism for the greet behavior) iii) QRIO kicks the ball and completes the plan.

The activation level traces for each behavior in a typical trial appear in figure 17 while the intentional levels are shown in figure 18. Figure 19 shows the experiment while it is being executed. As seen from the activation levels, each item of the plan is executed in sequential order mimicking the effect of a traditional deliberative controller directing the behavior of the robot. Note the large spike in activation level for the approach behavior D (figure 17) while QRIO is executing the plan. This occurs when the operator shouted and tried to gain QRIO's attention and showed QRIO the green ball. The greeting behavior's activation level increased greatly but the active maintenance of the intentional levels by the deliberative system prevents the high priority plan from being interrupted as shown by C in figure 18. In all trials, QRIO was able to execute the plan despite operator distraction.

6.3 Novel Behavior via Moderate Intentional Bias

In the third experiment, the same soccer scenario was used as in the previous experiment. In this experiment, however, only a moderate amount of intentional bias was introduced into the system when executing the plan. We hypothesize that moderate that intentional bias can support plan execution without subsuming the underlying reactive action selection mechanism. This will allow the robot to produce novel action sequences through the interaction of high-level plans and the action selection mechanism. Varying levels of intentional control will serve to make the robot's actions more natural and less scripted.

In the third experiment, the magnitude of the bias for each behavior in the search, approach, kick behavioral sequence was set to ($m = 75$). The plan is then executed as before, with the operator attempting to distract QRIO by calling out and then showing QRIO a ball in an effort to interact. If successful, the robot's behavior should differ from the high intentional bias case despite the fact that the same behaviors are being executed in each plan.

Graphs of the activation and intentional levels for a typical trial can be seen in figures 20 and 21 respectively, while figure 19 shows the trial. This time, when the operator tries to get QRIO's attention and shows him the green ball, the activation level exceeds that of the particular plan item, even with active maintenance (Figure 20, D). QRIO's desire to interact, driven by underlying action selection mechanism, causes the plan to be interrupted. Instead of continuing the plan, QRIO interacts with the user, and then resumes the plan from the point of interruption and finshes executing the soccer plan. In this example, the interplay between the low-level action selection mechanism and the high-level intentional control allowed QRIO to perform action sequences that were not explicitly planned. This allows QRIO to express a richer set of behavior then could be exhibited through purely deliberative control or control grounded solely in the reactive action-selection mechanism.

7 Conclusions and Future Work

This work presents the concept of intentional bias for deliberative control of reactive behavior. Through behavioral biasing, we show that goal-oriented behavior can be exhibited without overriding the underlying action selection mechanisms. In addition, through the active modulation of bias afforded by the intentional bus, the deliberative system is provided with a simple and uniform mechanism for interfacing with a reactive controller. In our implementation, we have investigated an additive mechanism for incorporating intentional bias into the reactive layer. In these experiments, it was shown intentional bias affords three types of deliberative control important for entertainment robots: plan execution, novel action production, and behavioral suppression.

Further research is necessary to determine if non-additive methods may be more effective in modifying the activation levels of the behaviors. In addition, it may be of value to investigate alternative means of calculating the

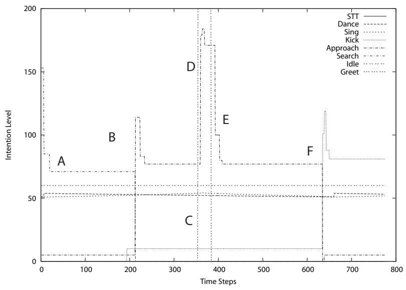

Fig. 20 Activation levels trace during the execution of a low intentional magnitude plan: A) Search behavior is active. B) Approach behavior becomes active. C) Releasing mechanism for greet behavior is found and its activation is increased. D) Intentional bus maintains appropriate intention levels but intentional magnitude is low enough to allow other behaviors to execute. E) Greeting behavior finishes and approach behavior becomes active again. F) Kick behavior becomes active.

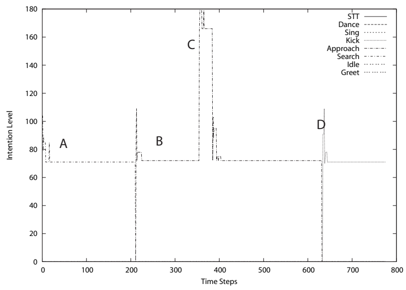

Fig. 21 Intention levels trace during the execution of a low intentional magnitude plan: A) Intentional bias is sent to search behavior. B) Intentional bias is stopped for search and started for approach behavior. C) Intentional bias is increased by intentional bus due to changes in behavior activation levels. D) Intentional bias to approach is stopped and bias to kick behavior begins.

(i) **(ii)** **(iii)**

Fig. 22 Depiction of low intention soccer plan. i) QRIO approaches the ball ii) The operator tries to distract QRIO with the green ball, the lower intentional bias of this plan allows the greeting behavior to run and QRIO greets the user and sits down. iii) QRIO gets back up, continues executing the plan, and kicks the ball.

magnitude of the bias for systems that do not have uniform activation level calculations over all behaviors.

Acknowledgements We would like to thank Ken-ichi Hidai, Kuniaki Noda, as well as the rest of the Sony Intelligence Dynamics Laboratories staff for their support in this work.

References

1. Andronache, V., Scheutz, M.: Contention scheduling: A viable action-selection mechanism of robotics? In: Proceedings of the 13th Midwest AI and Cognitive Science Conference, pp. 122–129. AAAI Press (2002)

2. Arkin, R.: Behavior-Based Robotics. MIT Press (1998)
3. Arkin, R., Fujita, M., Takagi, T., Hasegawa, R.: An ethological basis for human-robot interaction. Robotics and Autonomous Systems **42**(3-4), 191–201 (2003)
4. Blumberg, B.: Action-selection in hamsterdam: Lessons in ethology. In: Proceedings of the Third International Conference on the Simulation of Adaptive Behavior, pp. 108–117 (1994)
5. Bonasso, R., Kortenkamp, D.: An intelligent agent architecture in which to pursue robot learning. In: Working Notes: MCL-COLT '94 Robot Learning Workshop (1994)
6. Brazeal, C.: Sociable machines: Expresive social exchange between humans and robots. Ph.D. thesis, MIT (2000)
7. Burgard, W., Cremers, A., Fox, D., Hähnel, D., Lakemeyer, G., Schulz, D., Steiner, V., Thrun, S.: Interactive museum tour-guide robot. In: Proceedings of AAAI-98, pp. 11–18 (1998)
8. Carpenter, P., Riley, P., Veloso, M., Kaminka, G.: Integration of advice in an action selection architecture. In: Proceedings of the RoboCup2002 Symposium, pp. 195–205 (2002)
9. Cheranova, S., Arkin, R.: From deliberative to routine behaviors: A cognitively-inspired action selection mechanism for routine behavior capture. Adaptive Behavior **15**(2), 199–216 (2007)
10. Cooper, R., Shallice, T.: Modeling the selection of routine action: Exploring the criticality of parameter values. In: Proceedings of the 19th Annual Conference of the Cognitive Science Society, pp. 130–135 (1997)
11. Cooper, R., Shallice, T.: Contention scheduling and the control of routine activities. Cognitive Neuropsychology **17**(4), 297–338 (2000)
12. Fujita, M., Kuroki, Y., Ishida, T., Doi, T.: Autonomous behavior control architecture of entertainment humanoid robot SDR-4X. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 960–967 (2003)
13. Garforth, J., Meehan, A.: Driven by novelty? Integrating executive attention and emotion for autonomous cognitive development. In: Proceedings of the Developmental Robotics AAAI Spring Symposium (2005)
14. Garforth, J., Meehan, A., McHale, S.: Attentional behavior in robotic agents. In: Proceedings of the International Workshop on Recent Advances in Mobile Robots (2000)
15. Gat, E.: Artificial Intelligence and Mobile Robots, chap. On three-layer architectures. MIT/AAAI Press (1997)
16. Glasspool, D.: The integration of control and behavior: Insights from neuroscience and AI. In: Proceedings of the How to Design a Functioning Mind Symposium at AISB-2000, pp. 77–84 (2000)
17. Hoshino, Y., Takagi, T., Profio, U.D., Fujita, M.: Behavior description and control using behavior module for personal robot. In: Proceedings of the 2004 International Conference on Robotics and Automation, pp. 4165–4171 (2004)
18. Lyons, D., Hendricks, A.: Planning for reactive robot behavior. In: Proceedings of the IEEE International Conference on Robotics and Automation, pp. 2675–2680 (1992)

19. Maes, P.: How to do the right thing. Connection Science Journal **1**(3), 291–323 (1989)
20. Mastrogiovanni, F., Sgorbissa, A., Zaccari, R.: A system for hierarchical planning in service mobile robots. In: Proceedings of IAS08 (2004)
21. Moshinka, L., Arkin, R.: On TAMEing robots. In: Proceedings IEEE Conference on Systems, Man, and Cybernetics, pp. 3949–3959 (2003)
22. Norman, D.A., Shallice, T.: Consciousness and Self Regulation: Advances in Theory and Research, vol. 4, chap. Attention to action: Willed and automatic control of behavior. Academic Press (1986)
23. Okada, K., Haneda, A., Nakai, H., Inaba, M., Inoue, H.: Environment manipulation planner for humanoid robots using task graph that generates action sequence. In: Proceedings of 2004 International Conference on Intelligent Robots and Systems, pp. 1174–1179 (2004)
24. Payton, D., Rosenblatt, J., Keirsey, D.: Plan guided reaction. IEEE Transactions on Systems, Man, and Cybernetics **20**(6), 1370–1382 (1990)
25. Rosenblatt, J.: DAMN: A distributed architecture for mobile navigation. Journal of Experimental and Theoretical Artificial Intelligence **9**(2), 339–360 (1997)
26. Schallice, T., Burgess, P.: The domain of supervisory processes and temporal organization of behavior. In: Philosophical Transactions of the Royal Society of London B, vol. 351, pp. 1405–1412 (1996)
27. Wagner, A., Arkin, R.: Internalized plans for communication-sensitive team behaviors. In: Proceedings of the International Conference on Intelligent Robotics and Systems, pp. 2480–2487 (2003)