Chapter 7

Attentional Systems

Chapter Contents

Attention is a key component of all higher-level reasoning. A simplistic view is that the mechanism of attention seems to turn parts of the brain on and off in order to focus on what is currently important and ignore other things, but that does not tell the whole story. Characteristics of such attention processes depend on instincts, experiences and learning, and the human's goals and motivations. Attention helps us to cope with the large amount of information that can be acquired with our sensory systems in a short amount of time. In a sense, attention is one method that has evolved to ensure that we can succeed in the face of information overload. It helps us cope with complexity. Attention "filters out" less useful information from our senses (it "selects" the useful information), and thereby tries to optimally allocate cognitive resources. It also helps manage the complexity of internal reasoning (e.g., problem solving) by allowing us to focus on different internal representations, subproblems, and abstractions (i.e., it "selects" what to focus on when we are reasoning). Identifying components of attention is in fact complicated, as it closely intermixes with what are often considered other types of cognitive functions (e.g., planning and learning).

Due to its fundamental role in cognition, attention affects each type of control function that we have already considered in this part. On the other hand, control functions can affect attention since they dictate the behavior of dynamical attentional focusing. We may plan what to attend to, and have specific "attentional control rules" for how to focus. We can learn that certain stimuli are important to attend to since they help us reach our goals, or that such stimuli may play a significant detrimental role in our survival. We may learn that other stimuli can be ignored (i.e., learn that attending to some stimuli has no value to meeting our objectives). Indeed, we may even learn strategies for improving our attentional capabilities (e.g., how to concentrate better). We will not treat integrated attention-learning-planning in detail in this book. Instead, we will focus on the principles of dynamic focusing of attention, and analyze how control strategies can be used in attentional processes. We only briefly discuss how attentional strategies can be used in engineering applications for control and automation.

7.1 Neuroscience and Psychology of Attention

Attention is the process of focusing or concentrating. Often, we think of a hierarchy involving, in order of higher to lower levels, consciousness, sleeping, awareness, and attentiveness (e.g., you cannot be highly attentive when you are unconscious or asleep). At different points in our day we may turn off our attentional system. At others times our attentional system may be quite actively switching focus among different types of sensory data. For example, it may at one time disengage from one focus, move, and then engage on another focus. (Sometimes this is called "vigilance".) Attentional processes in the human brain are implemented with neural networks, but we will not consider this here (however, Design Problem 7.5 does request that you study the simulation of connectionist models of attentional systems).

We have certain types of attentional capabilities with all of our senses. For vision, we can pay attention to the object that we are looking at (e.g., focusing on these words as you read, while ignoring other peripheral visual stimuli or sounds). For auditory sensing, we may learn how to ignore background noise so that we are not distracted by it (e.g., if you have lived by a railroad track for a long time you may find yourself not even noticing a periodic passing train). For taste or smell, you can focus your attention on a certain spice in a food to try to identify it. For touch, we often ignore certain tactile senses (e.g., if you are holding this book, just an instant ago you probably did not notice the feeling of touching the book because you were probably attending to comprehension of the writing).

Attention allows us to amplify some sensory signals or internal thought processes and attenuate others.

A classical example of characteristics of our attentional system is given by the so-called "cocktail party effect." If you are at a party and there are many small groups of people talking, you have the useful ability to ignore (attenuate) what everyone is saying except for one person. The intriguing aspect, however, is that the person you are attending to (amplifying their signal) is not necessarily the one who is right next to you and talking the loudest. You may be able to virtually ignore this person to listen in from a distance on a quieter conversation that you are interested in (i.e., you may "eavesdrop").

In one famous experiment on human attention, "event-related potentials" (ERP) are measured via sensors on the scalp of a man via detection of electromagnetic waves. A specific ERP signal is the so-called "auditory N1 potential." The average voltage response for this ERP to an auditory stimulus that is attended to is relatively large in magnitude compared to an auditory stimulus that is not attended to. Some signals in the brain are amplified due to attention, and attenuated via lack of attention.

7.1.1 Dynamically Changing Focus

In the context of vision it is useful to think of our focus of attention as a type of "spotlight." This spotlight may coincide with where our eyes are focused ("overt" attention) or it may be that our eyes are focused at one point, and we attend to (shine our attentional spotlight) a different point ("covert" attention). Generally, we think of the spotlight as illuminating (amplifying) a region of sensory input. The dark region outside the spotlight is the region you are not attending to, and that visual sensory data are significantly attenuated.

There are two general types of control of attention, split according to what dictates the changes in attentional focus (i.e., what controls the dynamics of how the spot light moves). These are as follows:

Dynamic refocusing of attention can be driven by sensory data or explicit cognitive control.

- *Goal-driven (often "voluntary") attention reorientation:* Executive functions in the brain may reorient the focus of attention. This is thought of as a "top-down" refocusing that may be based on our problem-solving strategy and goals. For example, if you are reading and you decide to review a topic, you may go to the index of the book, find a key word, then go to another page and shift your focus of attention to another topic.

Typically, goal-driven reorientation of attention is slower and somewhat less "potent" than the stimulus-driven reorientation of attention that we discuss next.

- *Stimulus-driven (often "involuntary") attention reorientation:* Sensory signals can control the focus of attention in a "bottom-up" fashion. For example, we seem to have an instinct to pay attention to certain visual stimuli such as an object that is moving on a trajectory toward us, a bright flash of light (e.g., a fire), or blood (with evolutionary forces likely at work). Sensory inputs can achieve an *automatic* reorienting of attention, and often stimulus-driven attention reorientation is faster and more potent than goal-driven reorientation. For instance, if while reading this book, suddenly someone calls your name, yells in your ear, or your shirt catches on fire, it is likely that your attention will be diverted from this topic, no matter how interesting it is! Note, however, that if we repeatedly receive some external cue, and that cue does not indicate danger and we are not interested in it, we can typically learn to ignore it (i.e., learn not to allow sensory signals to reorient our attention). Hence, learning can play a key role in how our attentional dynamics operate.

Often, the two above methods to reorient attention are combined, or are interlaced over time. Clearly, both are influenced by knowledge acquired, and our instincts that have been established via evolution.

7.1.2 Multistage Processing: Filtering, Selection, and Resource Allocation

A functional model of the multistage attention process is given in Figure 7.1. There are sensory inputs that are "registered" (e.g., the receptor neurons detect sensory stimuli), then information is passed to the perceptual analysis and semantic encoding and analysis stages, where objects are recognized and processed for meaning. Information is then passed to executive functions, decision-making, memory, planning, etc. At the same time, there is feedback from executive functions that indicate what should be focused on (e.g., for voluntary control of attention).

Attention involves filtering out (discarding) some information.

There is evidence that at times, very early in the sensory processing process, there is selection of which stimuli are important, and which can be "filtered out." Evidence shows that in some situations this can be done before perceptual analysis or sensory encoding and analysis. For example, it seems that we have instinctual rules about certain types of stimuli that result in stimulus-driven reorientation of attention. It is this type of attentional control process that is involved in "early selection." On the other hand, "late selection" occurs in some situations, where more abstract analysis and processing of sensory signals (e.g., semantic encoding where meaning is determined) is conducted in order to reorient attention. For example, in some cases there might be some processing that determines whether the stimulus should gain full access to awareness, be

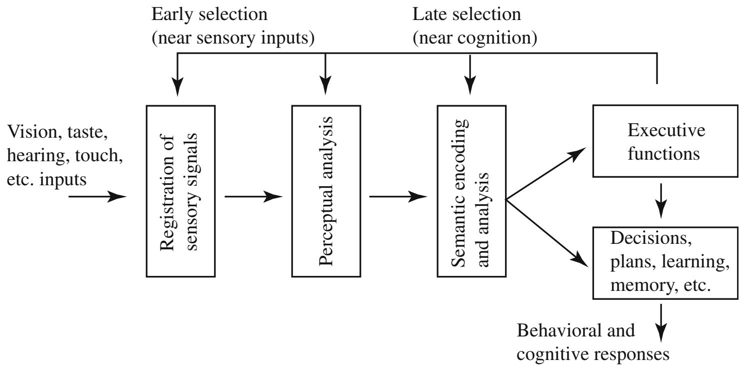

Figure 7.1: Multistage attention process.

encoded in memory, or result in some response. This type of processing may result in a goal-driven reorientation of attention.

Clearly, attention is a multistage process with feedback control paths. There is a cascaded filtering process that occurs where the most important information is focused on ("selected"), and less important information is ignored. Clearly, such a process is essential for high level cognitive functioning in humans. We have a finite amount of memory and processing power in our brain, and this naturally leads to "bottlenecks" in information processing. Attention allows us to allocate our cognitive resources to help us meet our goals. Hence, a key aspect of attention is the strategies used to allocate cognitive resources, especially in an "optimal" manner.

7.2 Dynamics of Attention: Search and Optimization Perspective

Here, we briefly discuss how to represent some of the underlying mechanisms of the dynamic focusing of attention as a search and optimization process. We will only focus on the dynamics of tracking objects in an attentional focus, how switching occurs from focusing on one object to another, and then the fine-tuning of the focus of attention after refocusing and during dynamic movement of an object. Our "model" is only based on the brief description in Section 7.1 of the psychology of attention, not neurophysiological studies, biophysics, or any of the other relevant underlying science. Hence, this is certainly of limited or possibly no value from a scientific perspective. Then, why provide such a model? First, the objective is to provide more detailed insight into the explanation of attention in Section 7.1. Second, we do not necessarily need a good model for the development of control and automation systems. The objective is to get the reader to think about dynamically focusing on information and hence ignoring other information. This is an essential feature of a complex automation problem where there is potential for "information overload" for the decision-making system, and hence the need to focus on the most important information.

We assume that there is a search component that finds objects in the "field of view" of the sensor and there is an optimization strategy at work that chooses the highest priority object and tracks it as it moves through the field of view. We do not focus on issues of the difference between where the sensor is directed versus where the focus of attention lies. Consider the model of attentional focusing provided in Figure 7.2. There, we have sensory data entering from the left into a block that processes these data to recognize objects, which we label with $i = 1, 2, \ldots, N$ (we assume a finite number of objects are in the field of view). For convenience, we assume in our discussion that the data are sensed about objects in an (x, y) plane. Next, the objects are prioritized by assigning a number $p_i > 0$, $i = 1, 2, \ldots, N$, where an object that is more important to focus on is given a higher value of p_i. Next, we assume that an "attention map" $J_a(x, y, t)$ is adjusted to represent the object positions and priorities at time t. Then, the priorities and attention map are input to a module that controls the focus point (i.e., where the focus is located in the (x, y) plane). To achieve control, it first compares the priorities to each other and picks the object $i^*(t)$ to focus on at time t that has the highest priority. That is, it lets

Dynamic focusing of attention can be modeled as optimization of a time-varying cost function.

$$i^*(t) = \arg \max_{i=1,2,\ldots,N} \{p_i\}$$

(arg max is simply the notation for finding the *index* of the priority that has the maximum value). Next, to pick the focus point, which we call (x_a, y_a), it considers which object should be focused on, where the current focus is relative to that object, and updates the focus point. It is assumed that it cannot move the focus point arbitrarily fast when it is trying to maintain focus on a particular object (e.g., as it moves across the plane), but that it can switch focus from one object to another very fast.

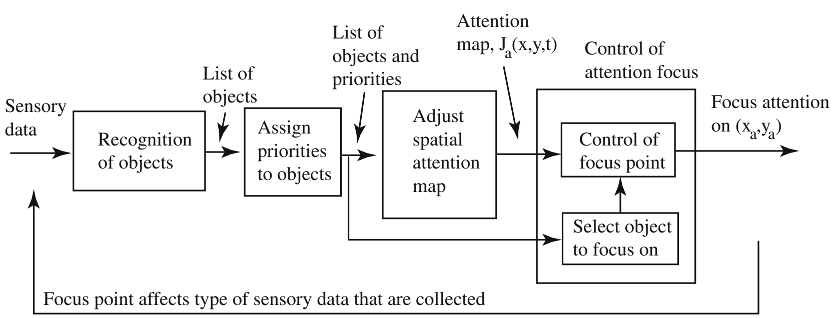

Figure 7.2: Functions involved in dynamically focusing attention.

7.2.1 Attentional Map

The key to our model of the dynamic focusing of attention lies in the definition of the attentional map $J_a(x, y, t)$. Here, we think of this map being generated internally (e.g., via pattern recognizer/semantic analysis), and assume that it is being used to indicate where it is important to focus on. In particular, we will define it as a continuous surface with $J_a(x, y) \in [-1, 0]$, where the point

$$(x_a, y_a) \in \{(x^*, y^*) : J_a(x^*, y^*) \leq J_a(x, y) \ \forall x, y\}$$

is a minimum point on the surface (note that there could be more than one such point, representing the possibility of a demand for split attention between equal priority points).

An example attentional map is shown in Figure 7.3. Here, we show an example attentional map that represents that there are two objects in the field of view: one that is high priority (the deeper valley) and one that is not as important (the shallow valley). The point that we want to focus on is the one defined by the point where the minimum is achieved on this map; that is, where the highest priority object is located.

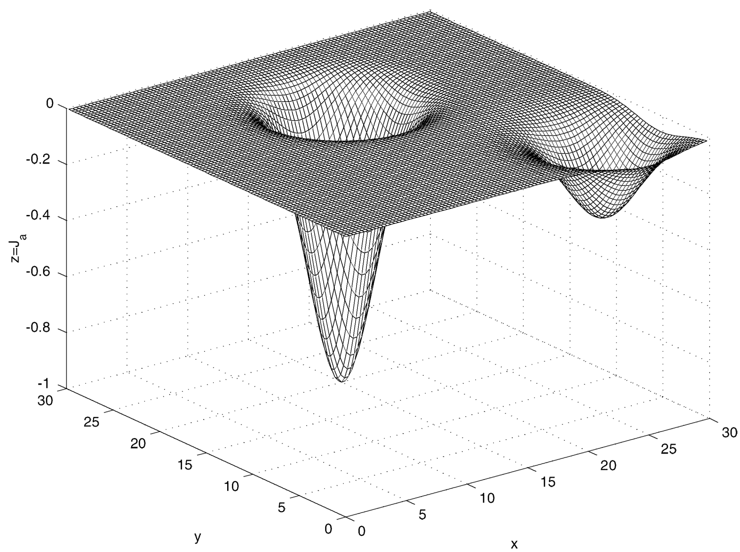

Figure 7.3: Example attentional map.

It is important to note that the map shown in Figure 7.3 is not static. It changes in several ways. First, if the objects move, the valleys move dynamically about the field of view. Also, if the field of view of the sensor is changed, the positions of the valleys change. The shape of the valleys may change (e.g., the widths of the valleys) depending on object positions. Moreover, the p_i priorities of the objects may change, which would dynamically change the depths of the valleys. For example, it may be possible that as relative positions of the objects changes, the shallow valley gets deeper and the deeper valley more shallow. If this happens, we would want the attentional focus to change from one object to another. If the valleys move, we want the attentional system to "track" the object that is of highest priority. Finally, note that as the field of view changes, and new objects appear and some disappear, it is possible that the number of valleys changes dynamically over time.

7.2.2 Optimization/Search Process for Focusing

It should be clear that in order to implement an attentional strategy using the attentional map, one could take an optimization/search perspective that has the following two components:

- *High priority object tracking:* Suppose that the current focus of attention (x_a, y_a) is located at the global minimum of the attentional map $J_a(x, y, t)$ (all other points on the J_a map are strictly above this point). Suppose that the field of view is constant, that objects do not leave the field of view, and that priorities of objects in the field of view stay constant. Suppose, however, that all the objects are moving and that some cognitive process keeps the attentional map up to date by dynamically adjusting the map. This will result in the centers of the valleys moving about the field of view dynamically. How does the attentional system work with the attentional map in order to maintain focus directly on the highest priority target? We could use a hill-climbing algorithm to continually climb down the attentional map at each step (e.g., it could move the focus of attention point in steps according to how someone would walk down a hill, moving in directions at each step toward the most significant decrease in the attentional map). Then, if the map does not move too fast, and the hill-climbing algorithm can keep up, it will tend to keep the focus of attention near the center of the valley that corresponds to the highest priority object. As the object moves about the field of view, the algorithm will tend to track the object.

- *Changing focus:* Next, suppose that the objects move about the field of view, and their priorities change dynamically. In this situation, the attentional tracking algorithm may track the highest priority object for a period of time, but its priority may decrease, and the priority of another object may increase. At the point where the global minimum of the attention function changes to correspond to the object with increasing priority, it should be the case that the strategy can switch focus from one object to another. How can this be achieved? Well, if the minimum points are known, switching is easy via a simple monitoring of the values of the minimum points of the attention function, ranking those values, and choosing to focus on the smallest one (a simple type of optimization). If those minimum points are not known, then one would need some type of "global" optimization procedure to determine when to switch. One approach would be to have N tracking algorithms of the type described above, and simply select for focusing the one that achieves the lowest value. There is some evidence that an analogous strategy is used in some cases in some biological attentional systems.

In summary, we see that one way to view the attentional process is as an optimization process for a cost function that is time-varying. Such an optimization problem can be very difficult to solve, but ideas from the optimization methods discussed in Part III and Part V provide many approaches to the problem.

7.3 Attentional Strategies for Multiple Predators and Prey

Consider an organism that is in some environment with multiple predators, and it is trying to attend to all of them to maintain as accurate a picture of its environment as possible in order that it can defend itself. Moreover, we assume that in the same environment, there are multiple prey that the organism would like to pay attention to in case it decides to pursue one of these to kill and eat. How should the organism dynamically focus its attention on the predators and prey to ensure its success in foraging and surviving? In this section we will model such a problem and introduce a variety of attentional ("scheduling") strategies for focusing attention. Hence, we think of needing to schedule our cognitive resources in order to maintain an accurate view of the environment. We will simulate the strategies and discuss issues in their design.

7.3.1 Cognitive Resource Allocation Model

We will assume that there is a recognizer for predators and prey that provides information to our attention strategy, so that it simply needs to decide what to focus on (cognitively process). The focus here is on the selection process that can be occurring in either early or late selection, or both. The key is that there is a "limited channel" or one resource that must be shared, and the attention strategy must decide how it is shared. We ignore issues of the possible differences in where the organism's sensor is pointed (e.g., where its vision is directed), versus where the center of the focus of attention is.

Quantifying Length of Time Predators/Prey Are Ignored

Suppose that we assume that the number of predators and prey is constant and that we number them and denote the set of predators and prey as

$$P = \{1, 2, \ldots, N\}$$

Let t denote time. Let

$$T_i(t), i \in P, t \geq 0$$

denote the *last time at which predator/prey i was detected* (later you will see that this is defined by the instant t' when, by focusing on predator/prey i, we get $T_i(t') = 0$). By "detected" we mean that the organism has focused its attention on the predator or prey, and has identified it and its characteristics (e.g., its position).

Generally, we will view the attentional strategies as "controllers" that take as inputs the $T_i(t)$, $i \in P$, and choose which predator/prey to focus on next. This is shown in Figure 7.4. We will assume that there is a cognitive tracking mechanism that is trying to estimate where predators/prey are moving, and that it has a certain level of accuracy in achieving this task. We will not require perfect accuracy in tracking multiple predators/prey; we will allow them to be "lost" for a period of time. Loss of tracking could result from predators or prey hiding (e.g., behind a tree), due to the sensor having only a limited "field of view," or from possible additional (but finite) time required to reacquire tracking when attentional focus is shifted. We will discuss how we model such issues in a moment.

An organism seeks to schedule its cognitive resources over time to enhance its chance of survival.

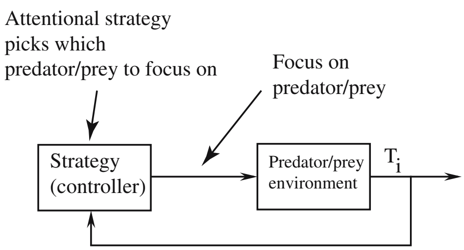

Figure 7.4: Attentional strategy viewed as a controller.

Suppose that initially

$$T_i(0) = 0, i \in P$$

so that we act as though initially we had simultaneously detected all the predators/prey, which is clearly physically impossible. Note, however, that this is a good initialization considering the fact that our attentional strategies will make decisions about which predator/prey to focus on based on the sizes of the $T_i(t)$, $i \in P$ (i.e., based on how long they have been ignored). Basically, for many strategies this initialization indicates that at $t = 0$, there is no priority to seek one predator/prey rather than any other one. For many strategies, an initialization with $T_i(0) > T_{i'}(0)$ for $i \neq i'$ would indicate an initial preference to first focus on the i^{th} predator/prey over predator/prey i'.

Note that if the organism was not actively engaged in paying attention to its environment (e.g., it was sleeping or doing something else), then clearly

$$T_i(t) \to \infty, i \in P, t \to \infty$$

since it will never detect a predator/prey. The goal of the attentional strategy is to try to avoid $T_i(t) \to \infty$ for any $i \in P$ and indeed it will try to keep the $T_i(t)$

values *as small as possible* since this represents that the organism has recently detected each predator/prey and hence has good information about the predators/prey. It is assumed that each predator/prey will *persistently* periodically "appear" (i.e., not be occluded by some object, or lost due to poor cognitive tracking) so that there is a finite amount of time between predator/prey appearances to the attentional strategy; this is assumed since, if some predator/prey $i \in P$ only appears for a finite amount of time, and never appears again, then at some point it will clearly be impossible to detect it again so that $T_i(t) \to \infty$ as $t \to \infty$.

The organism wants to minimize the amount of time it ignores any predator/prey to ensure it has accurate information about its environment.

Environmental and Cognitive Delays Affecting Attentional Switching

Let $\delta(t) > 0$ denote a "processing delay" that may represent the delay from the environment (e.g., due to a predator being occluded for a brief period of time) and a "cognitive processing delay." The cognitive processing delay may be used to represent the amount of time that it takes for the organism to switch from paying attention to one predator/prey i to another predator/prey j, $j \neq i$. We will call this type of delay $\delta_{i,j}$ and assume it is a fixed known delay (if it were unknown but bounded, then the attentional strategies and analysis still hold). For convenience, we will assume that these attentional switching delays are all the same and will denote that value by $\delta_s = \delta_{i,j}$ for all $i, j \in P$.

The variable $\delta(t)$ may also incorporate delays in being able to detect a predator/prey. For instance, each predator/prey has a type of frequency of appearance that is driven by a variety of characteristics such as how effectively the prey can hide in the current environment, or how fast a predator can run. Suppose that for a known predator/prey type i, there is some bound δ^i on the amount of time that it would take for the organism to first realize that the predator/prey may be at some location, if that was the only predator/prey that the organism focused on (clearly, this would depend on the predator/prey appearance period). Getting the first indication of the presence of a predator or prey does not correspond to achieving a detection of a predator/prey. Suppose that $\delta_e(t)$ denotes the delay incurred by the organism in first getting an indication of the presence of a predator/prey, from the time that it gets switched to focus on that predator/prey. It could be that many characteristics contribute to this delay, including cognitive tracking mechanisms and environmental characteristics. Note that if we let

$$\bar{\delta} = \max_i \left\{ \delta^i \right\}$$

then $\delta_e(t) \leq \bar{\delta}$. Let

$$\delta(t) = \delta_s + \delta_e(t)$$

For convenience, we let δ denote a constant that is the least upper bound on $\delta(t)$ so that $\delta(t) \leq \delta$ (i.e., we simply remove the time index to denote the least upper bound on the variable).

To summarize, when the attentional strategy issues a command to focus on predator/prey i, there is a delay to switch to the attention to focus on it, and then there is an additional (time-varying) delay since the predator/prey may not have appeared. This additional delay is shorter than $\bar{\delta}$. After these two types of delays occur we assume that the organism knows that a predator/prey is where it is focusing (but we do *not* assume that the organism has identified all the characteristics of the predator/prey and hence, has not yet "detected" it).

Rate of Cognitive Processing

We will suppose that the organism may take additional time to detect a predator/prey that it has not detected for a long period of time. That is, we think of the organism as having successively more difficult times finding a predator/prey that it has not found for longer periods of time since it, in a sense, becomes "desynchronized" with that predator/prey and cannot easily determine when or where it will appear, or its other characteristics. To quantify this phenomenon, we will use parameters

$$a_i, i \in P$$

where $1/a_i$ represents a "rate" at which the organism cognitively processes information about predators/prey in order to detect them. These a_i parameters require further explanation. Consider the case where there is only one predator/prey ($N = 1$), named "predator/prey 1." Suppose that at some time t', the amount of time that has elapsed since the last time predator/prey 1 was detected is $T_1(t') > 0$ as shown in Figure 7.5.

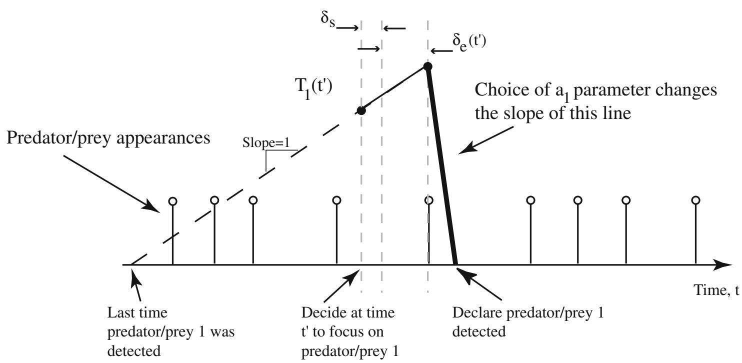

Figure 7.5: Illustration of timing of organism decision-making and predator/prey appearances (note that pulses represent the first times that predators/prey appear).

At time $t' + \delta_s$, the organism has switched its focus to predator/prey 1. So, starting at $t' + \delta_s$, the organism is looking for predator/prey 1 and before $t' + \delta_s + \delta^1$, we know that a predator/prey appearance will occur. Name the delay between achieving a switch in focus to the time where a predator/prey appearance is first found $\delta_e(t')$. Then, at time $t' + \delta_s + \delta_e(t')$, the organism initiates the completion of the "detection" of predator/prey 1 and the amount of time that it takes to do that is dictated by the a_1 parameter. (As you will see below, smaller values of a_1 correspond to it taking shorter amounts of time to fully detect the predator/prey.) We declare predator/prey 1 "detected" at the time at which T_1 is decreased to zero. Next, we need to further clarify the meaning of the a_i parameters by explaining how they produce the slope of the bold line in Figure 7.5 and hence quantify how long it takes to detect a predator/prey. Also, we need to explain how the organism chooses which predator/prey to focus on. To do this, we will introduce a specific attentional strategy and explain how to interpret the a_i, $i \in P$ parameters.

7.3.2 Focus on a Predator/Prey Ignored for the Longest Time

First, let D_{k_r} denote the time at which the attentional strategy chooses a predator/prey to focus on (i.e., it is the *decision time*), and suppose that $D_1 = 0$. An attentional strategy that focuses on the predator/prey that was ignored for the longest time makes choices of which predator/prey to focus on such that at D_{k_r}, the attentional strategy chooses to focus on predator/prey $i^*(k_r)$ such that

$$T_{i^*(k_r)}(D_{k_r}) \geq T_i(D_{k_r}), \forall i \in P \tag{7.1}$$

and focuses on it until it detects it. If there is more than one maximizer, then the attentional strategy will simply choose one of these at random.

Decision-Timing for Attentional Switches

First, notice that the actual time when focusing starts for predator/prey $i^*(k_r)$ occurs after some delay, and then it may take some additional (but finite time) for the predator/prey to appear ($\delta_e(D_{k_r}) \leq \bar{\delta}$), and still more time based on how long it has been since the predator/prey was last detected (i.e., the effect of the a_i). Note also that while the delays occur, the time since the last detection is still increasing. Hence, the times when the attentional strategy makes decisions are given by

$$D_{k_r+1} = D_{k_r} + \delta(D_{k_r}) + a_{i^*(k_r)} T_{i^*(k_r)}(D_{k_r}) + (D_{k_r+1} - D_{k_r})\, a_{i^*(k_r)} \tag{7.2}$$

Here, the next decision point D_{k_r+1} is the time when the detection of the last predator/prey that was focused on is detected and this formula gives the time D_{k_r+1} when the next decision will be made. The value of D_{k_r+1} is given by the sum of four terms. The first term is simply the last decision point D_{k_r}. The second term is the delay $\delta(D_{k_r})$ where

$$\delta(D_{k_r}) = \delta_s + \delta_e(D_{k_r})$$

Third, the term $a_{i^*(k_r)} T_{i^*(k_r)}(D_{k_r})$ is the amount of time it takes to detect predator/prey $i^*(k_r)$ that arises due to the fact that we have not detected it for some time. (Note the proportionality—if it has not been detected for a long time, then it will take more time to find it and this represents that predators/prey that have not been detected for a long time become more difficult to detect.) Finally, the fourth term quantifies that additional time is needed to detect the predator/prey simply because during the time that the cognitive processing for the predator/prey is occurring, even when it is focused on, the length of time since the last detection continues to increase (we do not consider a predator/prey $i^*(k_r)$ fully detected until $T_{i^*(k_r)}(D_{k_r+1}) = 0$).

Using simple algebra to rearrange Equation (7.2), we get

$$D_{k_r+1} = D_{k_r} + \frac{\delta(D_{k_r}) + a_{i^*(k_r)} T_{i^*(k_r)}(D_{k_r})}{1 - a_{i^*(k_r)}} \tag{7.3}$$

Notice that as expected, the delay δ directly influences the rate at which we can switch attentional focus. Also, this equation shows us that the length of time between decisions can be lengthened if a particular predator/prey has been ignored for too long due to the effects of the a_i parameters.

The Cognitive Capacity Constraint

In fact, using Equation (7.2), it is now possible to complete the explanation of Figure 7.5 and further explain how to interpret the a_i parameters. What is the effect of the a_i parameters on how fast a predator/prey is detected? Notice that we incur the delay $\delta(t)$, and from Figure 7.5 we see that the slope of the bold line dictates then how fast we achieve detection. What is the slope of the bold line in Figure 7.5? We use simple geometry to determine this. First, notice that the peak value

$$T_{i^*(k_r)}(D_{k_r} + \delta_s + \delta_e(D_{k_r})) = T_{i^*(k_r)}(D_{k_r}) + \delta_s + \delta_e(D_{k_r})$$

since the slope of the dashed line in Figure 7.5 is unity. Next, notice that Equation (7.3) gives the amount of time between the decision time D_{k_r} and time of detection D_{k_r+1} so that the slope of the bold line in Figure 7.5 is

$$-\left\{ \frac{T_{i^*(k_r)}(D_{k_r}) + \delta_s + \delta_e(D_{k_r})}{\frac{\delta_s + \delta_e(D_{k_r}) + a_{i^*(k_r)} T_{i^*(k_r)}(D_{k_r})}{1 - a_{i^*(k_r)}} - (\delta_s + \delta_e(D_{k_r}))} \right\}$$

which with some simple algebra reduces to

$$-\frac{(1 - a_{i^*(k_r)})}{a_{i^*(k_r)}} \tag{7.4}$$

In a moment you will see that it is necessary that $a_{i^*(k_r)} < 1$. Using this fact, Equation (7.4) indicates how fast detection occurs as shown in Figure 7.6 (i.e., how fast cognitive processing occurs). With small values of a_i (high values of $1/a_i$, the rate of processing by the organism in trying to detect) we get fast detection, and with larger ones we get slower detection. So, how do we interpret the a_i parameters? They are parameters used to model how difficult

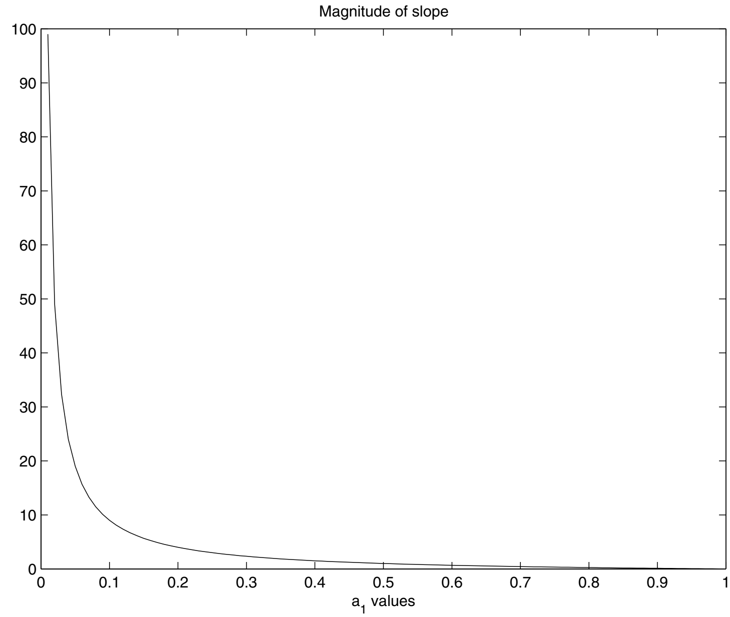

Figure 7.6: Magnitude of the slope of the bold line in Figure 7.5 for various values of a_1.

it is to detect a predator/prey, where if a predator/prey has not been detected for a long period of time, it can become more difficult to detect.

Clearly, it is *necessary* that the "capacity condition"

$$\rho = \sum_{i=1}^{N} a_i < 1 \tag{7.5}$$

be satisfied in order for *any* attentional strategy to ensure that the values of $T_i(t)$, $i \in P$, remain bounded. How should this capacity condition be interpreted? Intuitively, it says that it must be the case that even if the predators/prey can become more difficult to detect if they have not been detected for a long time, the organism must be able to operate "fast enough" to be able to find them. For instance, Equation (7.5) is satisfied if for each $i \in P$,

Cognitive capacity quantifies when an environment presents too large of an attentional load for an organism so that it will miss important information.

$$a_i < \frac{1}{N}$$

This shows us that as the number of predators/prey grows, it is possible that the cognitive capacity of the organism is overwhelmed and it is being given too much work, so that there is no way that it can keep up, so it will end up being the case that $T_i \to \infty$ for at least some $i \in P$ (or more than one i).

Equation (7.5) can be used to gain insight into the operation of attentional strategies by using the ideas in [418]. First, note that you can think of a_i as the amount of "load" (or the number of time units of "work") that is brought to the organism for the attentional task at each time instant by predator/prey i. Hence, if the organism is to succeed, on average the organism can only afford to spend a portion $(1-\rho)$ of its total time being idle. If you assume that the delay $\delta(t)$ is a constant δ, then each decision time when we switch from focusing on one predator/prey to another costs δ time units of idle time; hence, the average frequency of decision times is bounded above by

$$\frac{1-\rho}{\delta}$$

Now, if ρ is very close to 1 (representing an organism that is heavily loaded), $(1-\rho)\delta^{-1}$ is very small so the frequency of switching attentional focus between different predators/prey is low (which means that it can take a long time for the organism to find each predator/prey, so the organism will tend to have large $T_i(t)$ values and hence will not perform as well).

7.3.3 Additional Attentional Strategies

There are a wide variety of possible attentional strategies. Next, we consider one that is more general than the one of the previous subsection given in Equation (7.1), in the sense that at each decision point D_{k_r} it could make *exactly* the same decision as it did there, but could also make other choices.

Focus on a Predator/Prey Ignored More Than the Average One

The particular strategy is given by choosing the predator/prey to focus on that has been ignored more than the average time that all the predators/prey have been ignored. In particular, at D_{k_r}, the attentional strategy chooses to focus on predator/prey $i^*(k_r)$ such that

Attentional strategies are feedback controllers that dynamically refocus.

$$T_{i^*(k_r)}(D_{k_r}) \geq \frac{1}{N}\sum_{i=1}^{N} T_i(D_{k_r}) \tag{7.6}$$

and focuses on it until it detects it (in a similar way to the strategy of the last section). Note that Equation (7.3) also holds for this strategy, and that of course the capacity condition Equation (7.5) must hold.

Note that for this strategy, *any* predator/prey that has been ignored for more time than the average predator/prey has can be focused on. How does the strategy choose which particular predator/prey to focus on? One simple approach is to randomly choose one. However, more sophisticated criteria are possible. For instance, it could try to optimize some other system quantity, or it may use Equation (7.6) to provide a set of possible predators/prey to choose and then use "predator/prey priorities" (some indication of which predator/prey is most important) to choose the one to focus on. In the simulations of the next section, when we study this strategy, we will assume that predator/prey i has priority i and higher values of i correspond to higher priorities.

Focus on a Predator/Prey That May be Most Difficult to Find

An attentional strategy that focuses on the most difficult to find predator/prey makes choices of which predator/prey to focus on such that at D_{k_r}, the attentional strategy chooses to focus on predator/prey $i^*(k_r)$ such that

$$a_{i^*(k_r)} T_{i^*(k_r)}(D_{k_r}) \geq a_i T_i(D_{k_r}), \forall i \in P \tag{7.7}$$

and focuses on it until it detects it. If there is more than one maximizer, then the attentional strategy will simply choose one of these at random.

Clearly, this is similar to the attentional strategy that focuses on the predator/prey that has been ignored for the longest time that was given in Equation (7.1). Here, however, we have the scalings by the a_i parameters and this changes the attentional strategy. Intuitively, since a_i is the amount of "load," you can think of this attentional strategy as choosing the the most difficult one to find predator/prey to focus on.

Focus on a Predator/Prey Expected to be Most Difficult to Detect

The strategy to be developed next is motivated by the above strategy and is modeled after the one in [418] that has been found to be very effective in a different class of resource allocation problems. Recall from our earlier analysis that if you pick predator/prey $i^*(k_r)$ to focus on,

$$T_{i^*(k_r)}(D_{k_r}) + \delta_s + \delta_e(D_{k_r})$$

is the peak that $T_{i^*(k_r)}(D_{k_r})$ reaches before the predator/prey is detected. Note that in general we do not know $\delta_e(D_{k_r})$ since it depends on how the organism decision times are aligned with the predator/prey appearance times. A known bound, however, on the peak value is given by

$$T_{i^*(k_r)}(D_{k_r}) + \delta_s + \delta_e(D_{k_r}) \leq T_{i^*(k_r)}(D_{k_r}) + \delta_s + \delta^{i^*(k_r)}$$

Hence, predators/prey with larger δ^i values (i.e., ones with possibly lower frequency appearances) can be considered on average more difficult to detect in this framework. Also, the a_i parameters, which model another characteristic of the difficulty of predator/prey detection, will also affect how soon detection can occur.

Consider choosing predator/prey $i^*(k_r)$ to focus on at time D_{k_r} if

$$i^*(k_r) = \arg\max_i \left\{ w_i \left(\frac{T_i(D_{k_r}) + \delta_s + \delta^i}{\frac{(1-a_i)}{a_i}} \right) \right\} \tag{7.8}$$

where $w_i > 0$, $i \in P$ are weighting factors. Notice that in this formula, the numerator is the bound on the peak value and the denominator is the magnitude of the slope of the bold line in Figure 7.5 given by Equation (7.4). Why divide by the slope in the above formula? If the slope is greater in magnitude (smaller a_i value), this corresponds to an easier-to-detect predator/prey and this will result in Equation (7.8) with a *reduced* emphasis on focusing on that predator/prey. Hence, the strategy picks the predator/prey to focus on that is *expected* to be the most difficult to detect in the sense that it estimates which predator/prey will take the longest time to detect and selects it (assuming $w_i = 1$ for all i). To see this geometrically, notice via Figure 7.5 that the numerator $T_i(D_{k_r}) + \delta_s + \delta^i$ in Equation (7.8) should be thought of as an estimate of where the peak occurs and we divide it by the slope; hence, this value is the length of time that elapses from the time that the peak occurs, until detection.

The weighting factors w_i can be chosen to force the organism to focus on some predators/prey more than others. Equal weighting would correspond to the choice of $w_i = 1$ for all $i \in P$. If $w_i >> w_j$, $i \neq j$, then Equation (7.8) will tend to choose i rather than j to focus on. This may be useful in some predator/prey environments since it provides a way to indicate which predator/prey should be focused on. Another possibility is to weight predators more than prey so that the organism always focuses on those more. While the weighting factors provide an opportunity to tune the strategy, there is no guarantee that this strategy will be better than any of the others introduced above according to typical performance measures. Generally, you would want to choose the weights so as to make the attentional strategy perform as successfully as possible (where you define what is meant by "successfully").

7.3.4 Attentional Strategies Based on Predator/Prey Priority

In the last subsection, we introduced two ways to incorporate priorities of predators/prey into scheduling strategies. First, in Equation (7.6) we used priority as a "secondary" selection mechanism to choose from the set of predators/prey that has been ignored longer than the average one. Second, in Equation (7.8) we introduced the weighting factors w_i which allow us to emphasize the processing of one predator/prey more than another (and this will be illustrated in the simulation examples in Section 7.4). In this subsection we will introduce yet another priority scheme, but one that integrates the consideration of predator/prey priorities so that predator/prey priority is neither a secondary consideration nor set by secondary weighting parameters that have loose connections with the predator/prey priorities.

To do this, we introduce a set of parameters $p_i > 0$, $p_i \in \Re$, $i \in P$, that represent the predator/prey priorities (larger values correspond to higher priorities). We allow the designer to take two different views of the priority parameters:

Attentional strategies can include information on which predators/prey are most important to pay attention to.

1. *Predator/prey environment information:* You can assume that the values of the parameters p_i, $i \in P$, are set a priori and remain constant throughout the activity (e.g., foraging) of the organism. Hence, you can view them as part of the a priori information about the predator/prey environment.

2. *Design parameters:* Alternatively, you may view the priority parameters as design parameters that can be tuned (e.g., via extensive simulations of the

predator/prey environment) before an organism engages in the attentional task.

How do we integrate the priority parameters into *each* of the strategies defined in the previous subsections? For example, how can we use them to modify the strategy in Equation (7.1) where we chose to focus on the predator/prey that was ignored the longest. Here, we simply *scale* T_i by p_i, $i \in P$ in each of the cases and then make all decisions based on the same formulas as above, but with T_i replaced by $p_i T_i$, $i \in P$. What is the effect of such a scaling? It serves to scale the lengths of times that the predators/prey have been ignored, with higher weights given to predators/prey with higher priorities. Thereby, it biases the attentional strategy toward higher priority predators/prey.

For such strategies to be stable, it is clearly necessary that we modify our capacity condition. With priorities, we require that

$$\rho_p = \sum_{i=1}^{N} p_i a_i < 1 \tag{7.9}$$

be satisfied to ensure that the values of $T_i(t)$, $i \in P$, remain bounded.

How does the scaling affect the behavior of the strategies? While it is clear that predators/prey $i \in P$ with T_i scaled by higher values of p_i will have $p_i T_i$ grow faster (the slope of the line representing the growth is p_i), the behavior is also affected by the range of values that you allow for the priorities. For instance, if you dictate that your priorities $p_i \in (0, 1]$, $i \in P$, then if you were given some a_i values that satisfied Equation (7.5), the p_i and a_i values would also satisfy Equation (7.9). Hence, if you use a proper range of values for the priority parameters, any strategy that satisfies the capacity condition without priorities will satisfy Equation (7.9). Note that there is really no reason why you cannot make the choice of $p_i \in (0, 1]$, $i \in P$, since the parameters are simply used to rank order the predators/prey. It is also interesting to note that if you repeat the analysis in Sections 7.3.1 and 7.3.2, the result in Equation (7.4) still holds (due to cancellations of the priority parameters in the algebra); hence, simulation of the class of priority strategies discussed here is quite similar to the earlier strategies.

To summarize, *you can embed the priority parameters into any of the above strategies.* For instance, Equation (7.1), when converted to a priority scheme using this approach, becomes one where the attentional strategy chooses to focus on predator/prey $i^*(k_r)$ such that

$$p_{i^*(k_r)} T_{i^*(k_r)}(D_{k_r}) \geq p_i T_i(D_{k_r}), \forall i \in P \tag{7.10}$$

In this way you can have a strategy that selects predators/prey based on both priorities and how long they have been ignored. The scheduling strategies in earlier subsections are modified in a similar manner. Finally, note that you can still use the two priority schemes we discussed earlier in conjunction with this priority scheme.

7.3.5 Viewpoint of Attention Scheduling as Online Optimization

Next, note that we can provide an interpretation of the above attentional strategies in terms of optimization. The key is to think of the attentional decision-making in terms of *optimizing* a cost function J_p, and that J_p is the result of a computation made in the scheduler (controller) so that it can make scheduling decisions. With this view, we have the following:

Attentional strategies make decisions that optimize some short-term performance measure in hopes of optimizing a long-term one.

- *Focus on a Predator/Prey Ignored for the Longest Time:* Here, for the strategy in Equation (7.1), we have

  $$J_p = -\max\{T_i(D_{k_r}) : i = 1, 2, \ldots, N\}$$

  and hence in trying to maximize J_p, we try to minimize the longest time that the organism ignores any predator/prey. In this way, the scheduler tries to focus on predators/prey so as to keep the values of $T_i(t)$ low so that the organism has good information about the predators/prey.

- *Focus on a Predator/Prey Ignored More Than the Average One:* Here, for the strategy in Equation (7.6), we have

  $$J_p = -\sum_{i=1}^{N} T_i(D_{k_r})$$

  and hence in trying to maximize J_p, we try to minimize the average time that the organism ignores any predator/prey (it attempts this even though there is not a single maximizer at each decision time). Again, the scheduler tries to focus on predators/prey so as to keep the values of $T_i(t)$ low so that the organism has good information about the predators/prey. Here, however, it makes decisions in a different manner since it tries to maximize a different J_p.

Using this same approach, it is simple to specify J_p measures for the other strategies we defined above. For instance, for the strategy in Equation (7.7), we have $J_p = -\max\{a_i T_i(D_{k_r}) : i = 1, 2, \ldots, N\}$ and hence, in trying to maximize J_p, we try to minimize the longest time that the organism ignores any predator/prey, but scaled by the "load" of the predator/prey. For Equation (7.8), our J_p would quantify the desire to keep the peaks of the $T_i(t)$ as low as possible (which may or may not result in a lower average delay). Clearly, if you embed a priority scheme via the priority parameters p_i, $i \in P$, the same concepts hold.

Note that the above J_p measures should not be thought of as measures of attentional success over the long term, but as instantaneous measures that are used to guide decisions about which predator/prey to focus on. Achieving an instantaneous optimization does not necessarily result in making optimal decisions to try to ensure that the organism gets the best information over the long term.

7.4 Design Example: Attentional Strategies

In this section, we will simulate the attentional strategies of the last section in order to provide insights into their operation. Moreover, we will discuss several issues in how to design attentional strategies.

7.4.1 Simulation Approach and Performance Measures

For convenience, we simulate the predator/prey environment and organism as a discrete-time system. We will use a sampling period of $T_s = 0.01$ and in all our simulations we will have $N = 4$ predators/prey. Each predator/prey will be characterized by a sequence of appearances, which we simply model as unity height signal at some sampling instant. When there is no appearance, the signal height is zero. For instance, for all our simulations below we will have the predator/prey appearance sequences shown in Figure 7.7. We use different frequencies of appearance for different predators/prey, but for simplicity we keep the appearance frequencies constant (for predators/prey $i = 1, 2, 3, 4$ we have them appear every 1, 1.1, 1.2, and 1.3 sec.).

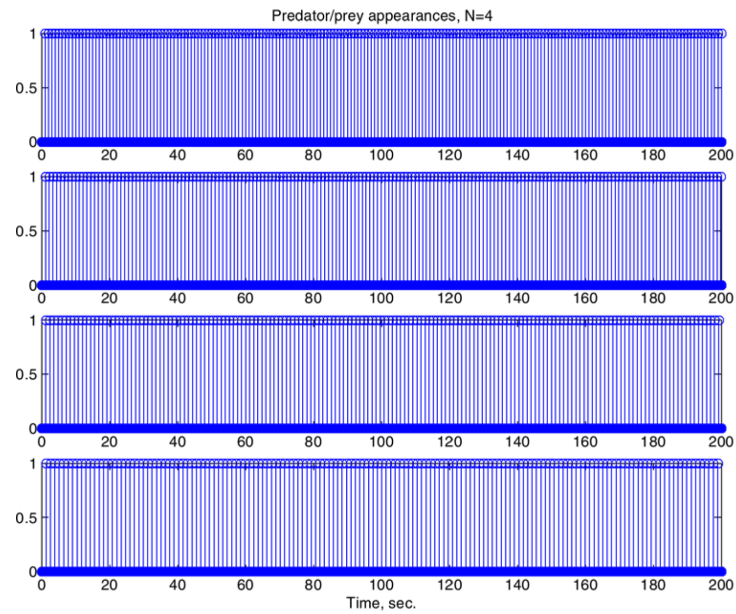

Figure 7.7: Predator/prey appearance sequences, for $N = 4$ predators/prey (predator/prey $i = 1$ is the top plot, $i = 2$ is the next one down, $i = 3$ is below that, and $i = 4$ is the bottom plot).

Suppose that we know that the bounds on the spacing between appearances are

$$\delta^1 = 1.05,\ \delta^2 = 1.15,\ \delta^3 = 1.25,\ \delta^4 = 1.35$$

Notice that these are simply bounds for periods given in Figure 7.7. We choose $\delta_s = 0.03$. To model detection difficulty, and in order to satisfy the capacity condition, we choose

$$a_1 = 0.1,\ a_2 = 0.2,\ a_3 = 0.3,\ a_4 = 0.1$$

This gives $\sum_{i=1}^{4} a_i = 0.7$, which represents that the organism will be quite busy in detecting predators/prey (lower values of this sum correspond to light loads).

There are several ways to measure performance of the attentional strategies. Here we will compute the average of the length of time since any predator/prey has been detected

$$\frac{1}{N}\sum_{i=1}^{N} T_i(k)$$

at each step k. We will also compute the time average of this quantity (i.e., the time average of the average values) and the maximum average value achieved over the entire simulation run. We will compute the maximum time that any predator/prey has been ignored at each time step k

$$\max_i \{T_i(k)\}$$

We will also compute the time average of this quantity (i.e., the time average of the maximum values) and the maximum of the maximum values achieved over the entire simulation run. In order to measure how well we have focused on higher priority predators/prey, we will use

$$\frac{1}{N}\sum_{k} i^*(k)$$

where $i^*(k)$ is the predator/prey chosen as step k. Clearly, higher values of this measure will correspond to the case where on average, higher priority predators/prey were focused on, in the case where we use i to both label the predators/prey and as a priority parameter.

7.4.2 Attentional Strategy Behavior: Focus on Longest Ignored

Here, we will illustrate the performance of the attentional strategy in Equation (7.1) that chooses the predator/prey to focus on that has not been detected for the longest period of time.

First, consider Figure 7.8 where the top plot shows $i^*(t)$, the predator/prey being focused on at each time. The plot below it shows $T_1(t)$, and the bottom plot shows $T_2(t)$. From the top plot it is interesting to note that the sequence of predators/prey that is focused on is: $1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4, \ldots$. But, the lengths of time that each is focused on is different, due to how the organism decision times happen to line up with the predator/prey appearances and due to the a_i values. Notice the periodic behavior of the $T_1(t)$ and $T_2(t)$ plots (due to the switching from focusing on one predator/prey to another). Figure 7.9 shows a similar plot, but for predators/prey 3 and 4. Notice that the periodic behavior of T_3 and T_4 is different from those shown in Figure 7.8. Ultimately, the pattern of the behavior of the $T_i(t)$ depends on the pattern of predator/prey pulses, the a_i values, the delay values, and how the predator/prey appearances align with the decision times.

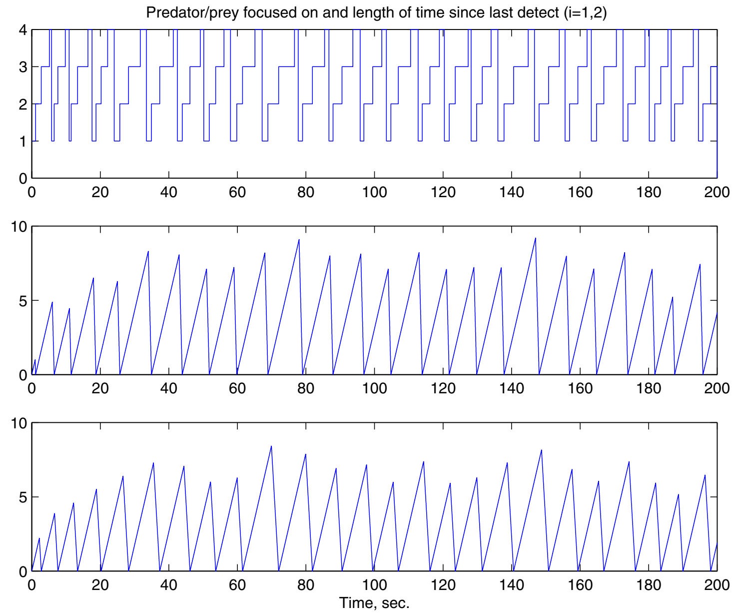

Figure 7.8: Attention scheduler decisions, and $T_i(t)$ for predators/prey 1 and 2.

Figure 7.10 shows a summary view of the dynamics of the attentional scheduling process. There, in the top plot, we also plot the average of the priorities of the predators/prey (assuming that priorities are defined by the i indices). The bottom plot shows the dynamics by showing all the $T_i(t)$ functions on one plot so that you can see the pattern of switching, and the maximum amount of time that the organism ignores any predator/prey. In Figure 7.11, we plot the performance measures of the average length of time since the last detection and maximum length of time since the last detection (and their average values as the straight lines).

Next, the program outputs some numeric values of the performance measures: (i) The time average of the priorities is 2.5670, (ii) the time average of the average values of the lengths of times waited is 3.4066, (iii) the maximum of the average values of the lengths of times waited is 5.7949, (iv) the time average of the maximum values of the lengths of times waited is 5.8297, and (v) the maximum of the maximum values of the lengths of times waited is 9.6199.

The time average of the average values is 3.4066, and this provides a good measure of scheduler performance. What does this value mean? It means that

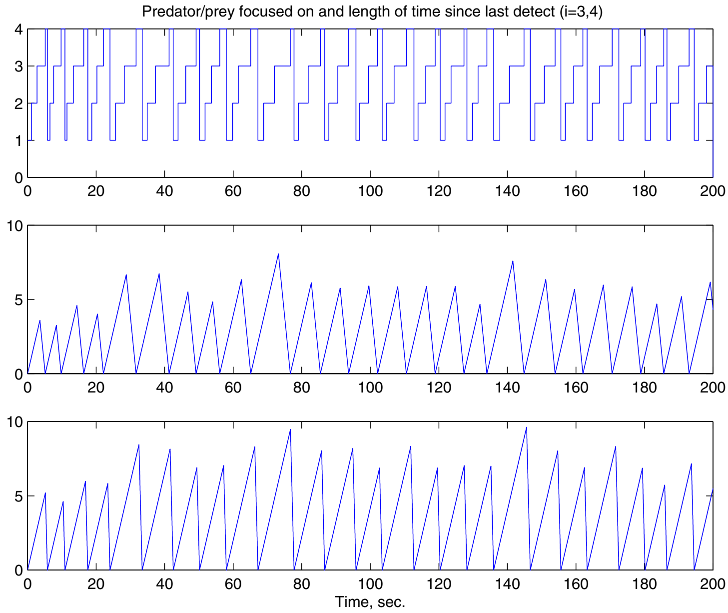

Figure 7.9: Scheduler decisions, and $T_i(t)$ for predators/prey 3 and 4.

on average, the organism detected each predator/prey every 3.4066 seconds. Is this good performance? Notice that the predator/prey appearances occurred every $1, 1.1, 1.2,$ and 1.3 seconds (predators/prey $i = 1, 2, 3, 4$ respectively). Considering the relative low rates of processing to detect the predators/prey, and the delays in switching and waiting for appearances, this appears to be reasonably good performance. Clearly, the performance could go up or down if the frequency or timing of the predators/prey appearances changed.

7.4.3 Effect of Focusing on Higher Priority Predators/Prey

Next, we use the strategy in Equation (7.6) that picks the predator/prey that has been ignored longer than the average one. For the set of predators/prey that has been ignored longer than the average one, we choose the one that has highest priority (i.e., predator/prey i with the greatest value of i). In this way, we study how priorities enter into attentional strategies by augmenting the strategy with a priority scheme. In this case, we get Figures 7.12 and 7.13. We see in Figure 7.12 that the sequence of predators/prey that is focused on is different from the previous strategy, and that the sequence is not periodic in the same way (e.g., it is not a simple $1, 2, 3, 4$ sequence). Also, we see that the average value of the priority of the predator/prey that is focused on is a bit higher, as we would expect. The bottom plot in Figure 7.12 shows quite a different behavior than the bottom plot in Figure 7.10; notice that here there is not an equal "balance" in focusing, since we see that the average values of

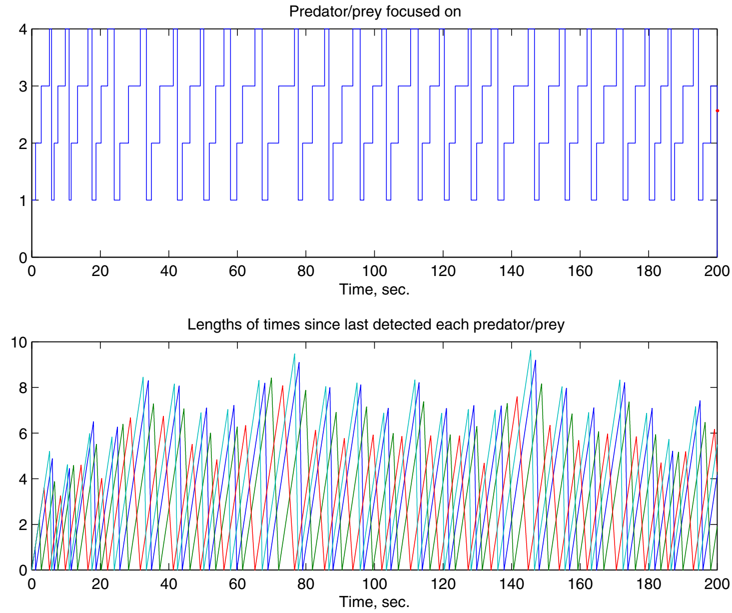

Figure 7.10: Attention scheduler decisions, and $T_i(t)$ for predators/prey $i = 1, 2, 3, 4$.

the $T_i(t)$ are quite different (e.g., see the occasional peaks). Next, notice that in Figure 7.13 we get poorer performance than that shown in Figure 7.11.

Frequent focusing on high priority predators/prey generally requires you to ignore others for longer periods of time.

To quantify the performance further, notice that the numeric performance measures are: (i) the time average of the priorities is 2.6896, (ii) the time average of the average values of the lengths of times waited is 3.8204, (iii) the maximum of the average values of the lengths of times waited is 6.4525, (iv) the time average of the maximum values of the lengths of times waited is 7.6574, and (v) the maximum of the maximum values of the lengths of times waited is 15.7399. This clearly shows that while we get *slightly* better focusing on higher priority predators/prey, we get poorer performance for all the other performance measures. We have paid a price in focusing on high priority predators/prey by ignoring other predators/prey for longer periods of time.

7.4.4 Tuning Attentional Strategy Parameters

As we saw with Equations (7.8) and (7.10), there are ways to define attentional strategies in terms of a set of parameters that specify how they make decisions (e.g., weights or priorities that modify J_p). For instance, we could specify the w_i weights such that there is a high emphasis on focusing on one predator or prey. To do this, you simply make one w_i value much larger than the others. This will result in frequent focusing on the corresponding predator/prey. Suppose that we are not concerned with predator/prey priority, or that all the predators/prey

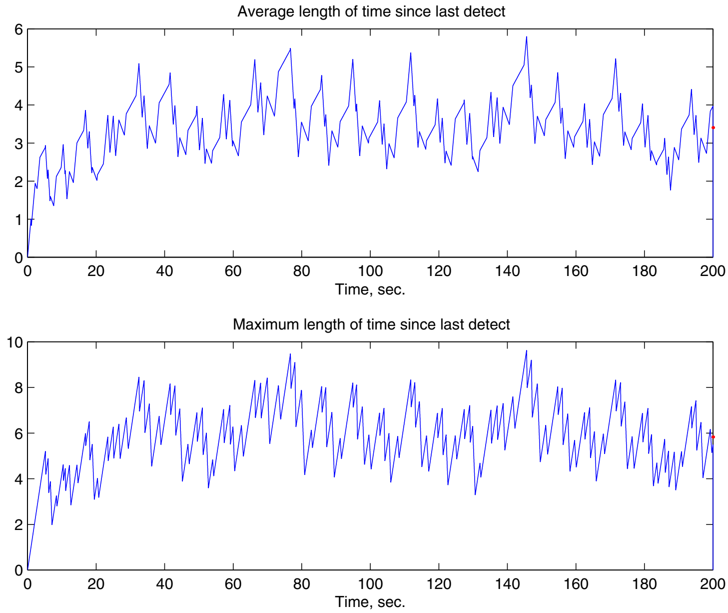

Figure 7.11: Performance measures (average and maximum times since last detection) and the time averages of their values.

have the same priority.

Can we tune the w_i values in Equation (7.8) in order to try to improve the performance measures? That is, can we use the parameters to simply try to improve performance, rather than emphasize focusing on a particular high priority predator/prey? The answer is yes, and to illustrate this, we ran a few simulations, tuning the w_i values with a focus on trying to minimize time average of the average values of the lengths of times waited. We obtained $w_1 = 4$, $w_2 = 2$, $w_3 = 1$, and $w_4 = 4$ and we get the performance in Figure 7.14. The tuning strategy used was to try a set of w_i values and look at the $T_i(t)$ plots. Then the value of w_i was increased a bit for the predator/prey that had higher peak values in order to try to make the strategy focus on that predator/prey more heavily.

The performance for this new set of w_i values is quantified via the following: (i) the time average of the priorities is 2.5802, (ii) the time average of the average values of the lengths of times waited is 3.2755, (iii) the maximum of the average values of the lengths of times waited is 5.3599, (iv) the time average of the maximum values of the lengths of times waited is 5.6423, and (v) the maximum of the maximum values of the lengths of times waited is 9.0899.

Notice that compared to the result in Section 7.4.2, we have tuned the w_i values to get a better value for time average of the average values of the lengths of times waited (there we obtained 3.4066). Is there further room to improve the performance of the scheduler? This seems likely, as the tuning process used

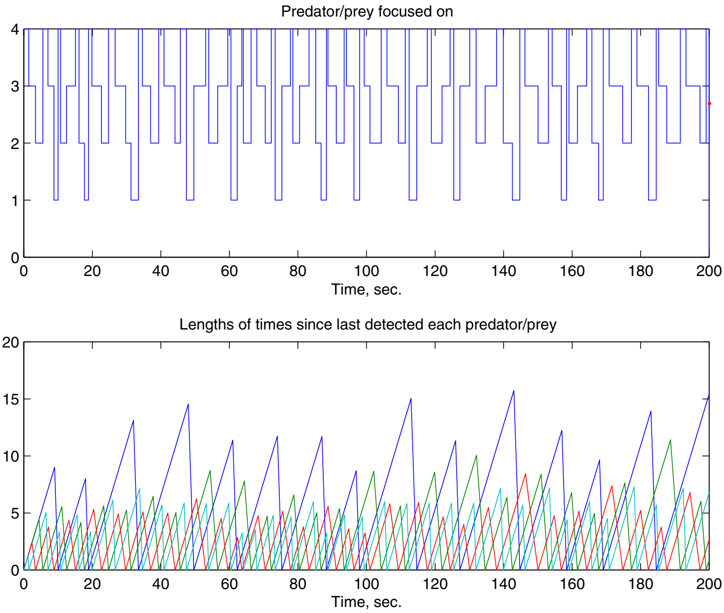

Figure 7.12: Attention scheduler decisions, and $T_i(t)$ for predators/prey $i = 1, 2, 3, 4$.

did not involve consideration of too many values of the parameters. It should be clear that the tuning problem can be quite difficult, especially if there are many predators/prey.

7.5 Stability Analysis of Attentional Strategies

In this section, the first three attentional strategies defined earlier will be proven to be stable, given that the capacity condition in Equation (7.5) holds. Stability of the strategy defined in Equation (7.8) can be studied using a similar proof procedure. Moreover, it is simple to extend the analysis below to the case where priority parameters are added as discussed in Section 7.3.4. At the end of this section, we will explain how to design a strategy that will stabilize *any* scheduling strategy, such as the ones that we will discuss in the next section.

7.5.1 Stability Properties of Attentional Strategies

We begin with the strategies defined in Equations (7.1) and (7.6).

Theorem 1: Assume that Equation (7.5) holds. The attentional strategies where the predator/prey that was ignored the longest time, or one that has been ignored longer than the average one, as defined in Equations (7.1) and (7.6),

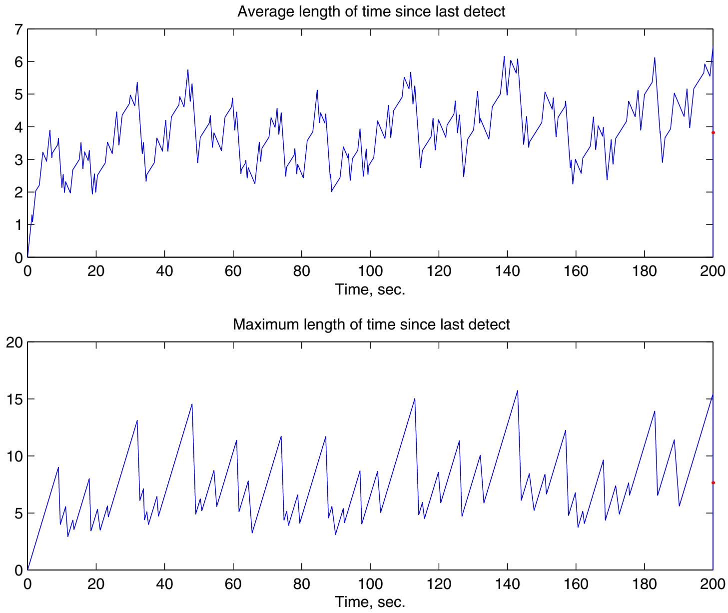

Figure 7.13: Performance measures (average and maximum times since last detection) and the time averages of their values.

have the following properties: They are *stable* in that

$$\sup_{t \geq 0}\{T_i(t)\} < B_i, i \in P$$

for some $B_i > 0$, $i \in P$ so that they will not ignore any predator/prey for too long. A specific bound on the ultimate longest time that the organism will ignore any predator/prey is given by

$$\lim_{t\to\infty} \sup \sum_{i=1}^{N} T_i(t) \leq \delta \left[\frac{\sum_{i=1}^{N} a_i}{\underline{a}} + \frac{\bar{a}N}{\underline{a}\left(1 - \sum_{i=1}^{N} a_i\right)} \max_i \left\{ \frac{-a_i + \sum_{i=1}^{N} a_i}{a_i} \right\} \right]$$

where $\underline{a} = \min_i\{a_i\}$ and $\bar{a} = \max_i\{a_i\}$.

Proof: Let

$$V(t) = \sum_{i=1}^{N} a_i T_i(t)$$

be a "Lyapunov-like" function (strictly speaking it is not a Lyapunov function because $T_i(t)$ is not the state of the system, e.g., due to the presence of the delays). You can think of $V(t)$ as the amount of work that the organism needs to do at time t in order to obtain perfect information about all the predators/prey.

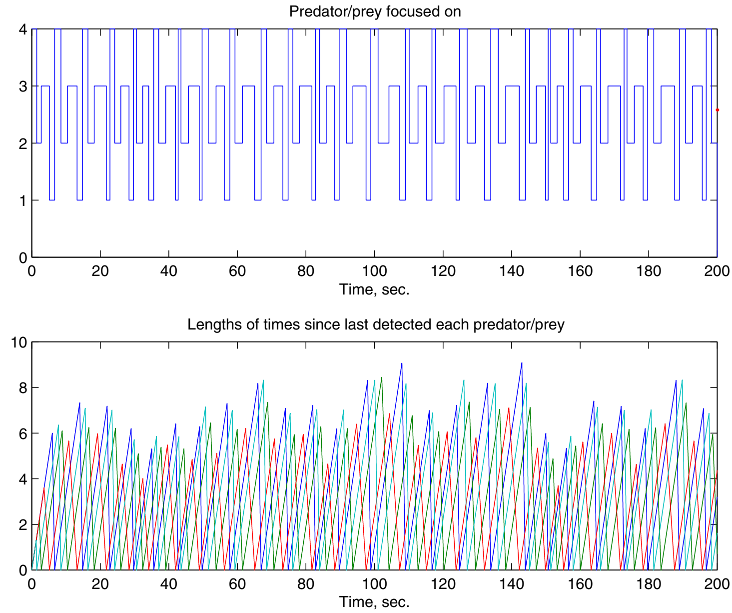

Figure 7.14: Attention scheduler decisions, and $T_i(t)$ for predators/prey $i = 1, 2, 3, 4$.

The proof to follow focuses on the strategy where the predator/prey is chosen that has been ignored longer than the average one; however, a special case of this is when the one that is ignored the longest is chosen at each decision point so the above bounds hold for that attentional strategy also.

Note that since $T_{i^*(k_r)}(D_{k_r+1}) = 0$ ($i^*(k_r)$ was the predator/prey that was just detected),

$$V(D_{k_r+1}) = \sum_{i=1}^{N} a_i T_i(D_{k_r+1}) = \sum_{i \neq i^*(k_r)}^{N} a_i T_i(D_{k_r+1})$$

Also,

$$\sum_{i \neq i^*(k_r)}^{N} a_i T_i(D_{k_r+1}) = \sum_{i \neq i^*(k_r)}^{N} a_i \left(T_i(D_{k_r}) + (D_{k_r+1} - D_{k_r})\right)$$

since when the organism is focusing on predator/prey $i^*(k_r)$, the amount of time that all other predators/prey are ignored increases by $(D_{k_r+1} - D_{k_r})$ for each i, $i \neq i^*(k_r)$. Rearrange this equation to obtain

$$V(D_{k_r+1}) = V(D_{k_r}) - a_{i^*(k_r)} T_{i^*(k_r)}(D_{k_r}) + (D_{k_r+1} - D_{k_r}) \sum_{i \neq i^*(k_r)}^{N} a_i \quad (7.11)$$

Now, use Equation (7.3) to obtain

$$V(D_{k_r+1}) \leq V(D_{k_r}) - \alpha(i^*(k_r))T_{i^*(k_r)}(D_{k_r}) + \beta(i^*(k_r)) \tag{7.12}$$

where

$$\alpha(i) = \frac{a_i\left(1 - \sum_{j=1}^{N} a_j\right)}{1 - a_i}$$

and

$$\beta(i) = \delta\frac{\left(-a_i + \sum_{j=1}^{N} a_j\right)}{1 - a_i}$$

Note that $\alpha(i) > 0$ and $\beta(i) > 0$ for all $i \in P$. To understand how Equation (7.12) is found, using Equation (7.3), note that since $\delta(D_{k_r}) \leq \delta$

$$\begin{aligned} V(D_{k_r+1}) &\leq V(D_{k_r}) - a_{i^*(k_r)}T_{i^*(k_r)}(D_{k_r}) \\ &+ \left[\sum_{j \neq i^*(k_r)}^{N} a_j\right](1 - a_{i^*(k_r)})^{-1}\left[\delta + a_{i^*(k_r)}T_{i^*(k_r)}(D_{k_r})\right] \end{aligned}$$

The term due to δ creates $\beta(i^*(k_r))$. For the remaining terms, besides $V(D_{k_r})$, by grouping we get

$$-a_{i^*(k_r)}\left(1 - \frac{\sum_{j \neq i^*(k_r)}^{N} a_j}{1 - a_{i^*(k_r)}}\right)T_{i^*(k_r)}(D_{k_r}) =$$

$$-a_{i^*(k_r)}\left(\frac{1 - \sum_{j=1}^{N} a_j}{1 - a_{i^*(k_r)}}\right)T_{i^*(k_r)}(D_{k_r})$$

and this is used to define $\alpha(i^*(k_r))$.

Next, notice that due to the definition of *either* attentional strategy

$$\alpha(i^*(k_r))T_{i^*(k_r)}(D_{k_r}) \geq \alpha(i^*(k_r))\frac{1}{N}\sum_{i=1}^{N} T_i(D_{k_r})$$

and due to the definition of $\bar{a}$,

$$\alpha(i^*(k_r))\frac{1}{N}\sum_{i=1}^{N} T_i(D_{k_r}) \geq \alpha(i^*(k_r))\frac{1}{N}\bar{a}^{-1}\sum_{i=1}^{N} a_i T_i(D_{k_r})$$

(since $\frac{a_i}{\bar{a}} \leq 1$). But notice that

$$\alpha(i^*(k_r))\frac{1}{N}\bar{a}^{-1}\sum_{i=1}^{N} a_i T_i(D_{k_r}) = \alpha(i^*(k_r))\frac{1}{N}\bar{a}^{-1}V(D_{k_r}) \tag{7.13}$$

Combine this with Equation (7.12) to get

$$V(D_{k_r+1}) \leq \left[1 - \bar{a}^{-1}N^{-1}\alpha(i^*(k_r))\right]V(D_{k_r}) + \beta(i^*(k_r)) \tag{7.14}$$

Subtract $\bar{a}N \max_i \frac{\beta(i)}{\alpha(i)}$ from both sides of Equation (7.14) and after a bit of algebra, you get

$$\begin{aligned} V(D_{k_r+1}) &- \bar{a}N \max_i \frac{\beta(i)}{\alpha(i)} \\ &\leq \left[V(D_{k_r}) - \bar{a}N \max_i \frac{\beta(i)}{\alpha(i)}\right] \left[1 - \bar{a}^{-1}N^{-1}\alpha(i^*(k_r))\right] \\ &+ \beta(i^*(k_r)) \left[1 - \frac{\alpha(i^*(k_r))}{\beta(i^*(k_r))} \max_i \frac{\beta(i)}{\alpha(i)}\right] \end{aligned}$$

Focus for a moment on the last term in this equation, and notice that

$$\beta(i^*(k_r)) \left[1 - \frac{\alpha(i^*(k_r))}{\beta(i^*(k_r))} \max_i \frac{\beta(i)}{\alpha(i)}\right] \leq 0$$

How do you get the last inequality? Note that $\beta(i) > 0$. If the $\max_i \frac{\beta(i)}{\alpha(i)}$ term is maximized at some particular value j, then clearly this value divided by any value considered in the maximization will be greater than or equal to 1.

Now, we have

$$\begin{aligned} &\left[V(D_{k_r+1}) - \bar{a}N \max_i \frac{\beta(i)}{\alpha(i)}\right] \\ \leq &\left[V(D_{k_r}) - \bar{a}N \max_i \frac{\beta(i)}{\alpha(i)}\right] \left[1 - \bar{a}^{-1}N^{-1}\alpha(i^*(k_r))\right] \end{aligned} \tag{7.15}$$

But, notice that the second term on the right-hand side of this equation

$$\begin{aligned} \left[1 - \bar{a}^{-1}N^{-1}\alpha(i^*(k_r))\right] &\leq 1 - \bar{a}^{-1}N^{-1} \min_i \left\{ \frac{a_i \left(1 - \sum_{j=1}^N a_j\right)}{1 - a_i} \right\} \\ &\leq 1 - \bar{a}^{-1}N^{-1} \left[\frac{\underline{a} \left(1 - \sum_{j=1}^N a_j\right)}{1 - \underline{a}} \right] \end{aligned}$$

Notice that

$$0 < \frac{\left(1 - \sum_{j=1}^N a_j\right)}{1 - \underline{a}} < 1$$

and

$$0 < \frac{\underline{a}}{\bar{a}} < 1$$

so that

$$0 < 1 - \bar{a}^{-1}N^{-1}\alpha(i^*(k_r)) < 1$$

which makes the mapping in Equation (7.15) contractive so that

$$\lim_{k_r \to \infty} \sup \left\{ V(D_{k_r}) - \bar{a}N \max_i \frac{\beta(i)}{\alpha(i)} \right\} = 0$$

But this (ultimate) bound is in terms of only the decision points D_{k_r}, $k_r = 1, 2, 3, \ldots$ Due to the delay δ, the $T_i(t)$ values can rise higher at times t not at the decision points. However, for $D_{k_r} \leq t \leq D_{k_r+1}$

$$V(t) \leq V(D_{k_r} + \delta)$$

But notice that

$$V(D_{k_r} + \delta) = \sum_{i=1}^{N} a_i T_i(D_{k_r} + \delta) = \sum_{i=1}^{N} a_i T_i(D_{k_r}) + \delta \sum_{i=1}^{N} a_i = V(D_{k_r}) + \delta \sum_{i=1}^{N} a_i$$

This gives us

$$\lim_{t\to\infty} \sup V(t) \leq \delta \sum_{i=1}^{N} a_i + \bar{a} N \max_i \frac{\beta(i)}{\alpha(i)}$$

and since

$$\lim_{t\to\infty} \sup \sum_{i=1}^{N} T_i(t) \leq \frac{1}{\underline{a}} \lim_{t\to\infty} \sup V(t)$$

we know

$$\lim_{t\to\infty} \sup \sum_{i=1}^{N} T_i(t) \leq \frac{\delta \sum_{i=1}^{N} a_i}{\underline{a}} + \frac{\bar{a} N}{\underline{a}} \max_i \frac{\delta \left(-a_i + \sum_{j=1}^{N} a_j\right)}{a_i \left(1 - \sum_{j=1}^{N} a_j\right)}$$

which gives the desired result. ■

Note that since the above bound for Theorem 1 may be conservative for some situations, it would be of interest to specify "tight" bounds since this would provide good guarantees for bounding the maximum time that a predator/prey is ignored.

Next, we will study the stability properties of the other strategy defined in the last section where we get a different bound on the maximum length of time that a predator/prey will be ignored by the organism. The analysis, is however, only slightly different and depends on the above proof.

Theorem 2: Assume that Equation (7.5) holds. The attentional strategies defined in Equation (7.7) have the following properties: It is *stable* in that

$$\sup_{t\geq 0} \{T_i(t)\} < B_i, i \in P$$

for some $B_i > 0$, $i \in P$ so that it will not ignore any predator/prey for too long. A specific bound on the ultimate longest time that the organism will ignore any predator/prey is given by

$$\begin{aligned} \lim_{t\to\infty} \sup V(t) &\leq \frac{\delta(N-1)}{1 - \sum_{i=1}^{N} a_i} \left(-\underline{a} + \sum_{i=1}^{N} a_i\right) + \delta \sum_{i=1}^{N} a_i \\ &\leq \delta \left[\sum_{i=1}^{N} a_i\right] \frac{N - \sum_{i=1}^{N} a_i}{1 - \sum_{i=1}^{N} a_i} \end{aligned}$$

where $\underline{a} = \min_i\{a_i\}$ and $\bar{a} \max_i\{a_i\}$.

Proof: Use the ideas from the proof for Theorem 1 and note that Equation (7.13) in this case is

$$\alpha(i^*(k_r))T_{i^*(k_r)}(D_{k_r}) = \frac{\alpha(i^*(k_r))}{a_{i^*(k_r)}} a_{i^*(k_r)} T_{i^*(k_r)}(D_{k_r})$$

$$\geq \quad \frac{\alpha(i^*(k_r))}{(N-1)a_{i^*(k_r)}} \sum_{i=1}^{N} a_i T_i(D_{k_r}) = \frac{\alpha(i^*(k_r))}{(N-1)a_{i^*(k_r)}} V(D_{k_r})$$

The $N-1$ factor appears, rather than N, for one i, $T_i = 0$. To complete the proof, simply take the same approach as in the remainder of the proof of Theorem 1, below Equation (7.13). ■

So, do these bounds give an indication of which of the three strategies is "best"? Unfortunately, they generally do not since the bounds can be conservative. It is for this reason that simulation analysis is generally needed to analyze the *performance* of particular strategies and determine which is best for a particular predator/prey environment.

7.5.2 Stabilizing Mechanism for Attentional Strategies

At times there is significant knowledge about the predator/prey environment and organism that is relevant to the design of attentional strategies. There is then a natural tendency to incorporate this information in the specification of the attentional strategy, often in the form of "scheduling heuristics." We will briefly discuss two such approaches in the next section. The problem with this approach, however, is that the resulting strategies may end up being somewhat nonstandard and there may be concerns about whether they will be stable.

Fortunately, the approach in [290] to specifying a "universal stabilizing mechanism" (USM) for any scheduling strategy actually holds for the attention scheduling problem. (Actually, the approach in [290] was developed for a fixed size delay and we have a time-varying but bounded delay; however, the proofs there can be directly extended to our case with no difficulty.) This mechanism can then be applied to *any* heuristically constructed attentional strategy, and you will be ensured that the overall strategy will be stable. In this section, we introduce the USM from [290]. In the next two sections, we introduce two types of schedulers that exploit predator/prey domain information to try to enhance scheduler performance, and which can be stabilized by the USM introduced here.

The USM allows the designer to focus on improving performance of the attentional strategy.

The key fact is that for the resource allocation problems we consider here, as long as the capacity condition is satisfied, it is possible to define a USM which, when used to supervise a scheduling strategy, will always result in stable operation. To define the USM, let Q denote a first-come first-serve (FCFS) priority queue for predators/prey that have been ignored for a long time. For instance, if predator/prey's T_i value becomes too large, we will have criteria for entering the queue at some time t'. If predators/prey i and j are in the priority queue, and j entered it before i,

$$Q = (\ldots, i, j, \ldots)$$

then when this priority queue is serviced, predator/prey j will be taken off the queue and focused on before predator/prey i (the "tail" of the queue is the first predator/prey listed after the "(" and the "head" of the queue is the predator/prey listed just before the ")" in the definition of Q above). We need some additional parameters to specify the USM. Let $L > 0$ be a large number satisfying

$$L > \frac{N\delta}{1-\rho}$$

where ρ is specified in Equation (7.5), N is the number of predators/prey, and δ is the bound on the maximum delay. Next, let

$$H_i > 0, \ i \in P$$

denote a set of parameters, the interpretation of which will become clear as we define the USM.

The USM is implemented by the following set of rules:

1. *Truncation rule:* The organism can process no predator/prey i longer than La_i time units. This means that if at time $t'+\delta_s+\delta_e(t')$ the organism starts to try to detect predator/prey i, then it can only try to detect it no longer than up to the time $t'+\delta_s+\delta_e(t')+La_i$. If detection occurs before that time, then the strategy acts as usual and selects another predator/prey to focus on. If, however, it has not yet detected predator/prey i by this time, it is forced to make a new decision (which could entail switching predators/prey). Note that if it does switch to another predator/prey, we assume that the progress it had made on predator/prey i is used, but that the time since it was last detected, T_i, begins to increase again.

2. *Rule for entering Q:* Predator/prey i enters the tail of the priority queue Q at time t if we have not just decided to focus on i or are currently focusing (cognitively processing) to detect i, and $T_i(t) > H_i$ (hence, the H_i are thresholds for when a predator/prey is placed in the queue).

3. *Predator/prey selection rule:* If Q is not empty when the organism has finished focusing on a predator/prey (either by achieving $T_i = 0$ or via rule 1 above), then the predator/prey at the head of the priority queue Q (i.e., FCFS) is chosen.

4. *Rule for leaving Q:* A predator/prey i leaves Q at the time $t'+\delta_s+\delta_e(t')$ where t' is the time point when predator/prey i was selected by rule 3.

5. *Rule for processing-time for a predator/prey from Q:* If predator/prey i from Q is chosen to be focused on, then beyond the time t' defined in rule 4, it is processed for La_i time units unless it is detected (i.e., $T_i = 0$) before this time elapses.

Notice that this is not simply another attentional strategy. It actually defines a "supervisor" for any attentional strategy (e.g., ones that exploit heuristic information from the problem domain) that ensures it will result in stable operation. If you have constructed a stable strategy, and you choose L and the H_i, $i \in P$, large enough, then the USM will *never* intervene. The USM simply truncates the processing of predators/prey that are not found fast enough, and via Q makes sure that predators/prey that have been ignored for too long will get attention. How do we pick L and the H_i parameters? If you pick $H_i = 0$, $i \in P$, then the USM simply enforces a type of FCFS strategy on predators/prey with $T_i > 0$, but it stops processing any predator/prey that is focused on too long. In this case, the USM always intervenes. As you increase the size of L and the parameters H_i, $i \in P$, the USM intervenes less frequently.

What is the value of the USM? In a sense, it frees the designer of attentional strategies from being concerned about the stability of the myriad possible attentional strategies (but of course, the stability analysis of Section 7.5 is still useful, particularly if the analysis helps to clarify how to design the strategy to achieve high performance operation). You can adopt a design philosophy where you construct a very complicated attentional strategy, possibly exploiting heuristic ideas about how to achieve the best performance. Then, you can augment such strategies with the USM and be assured that you will obtain stable operation. Essentially, the USM allows the designer to focus on the design of attentional strategies to improve attention scheduling *performance*. To illustrate this point, in the next section we will briefly discuss the design of two heuristic strategies, ones based on our intuitions about the problem domain.

7.5.3 Planning and Attention

In this section, we discuss two ways to use planning concepts from Chapter 6 in attentional strategies. Intuitively, this should make sense. We can plan how to pay attention to a set of predators and prey if we have some idea of how the environment might behave, and if we consider alternative predators/prey to focus on based on predictions about how they might behave. We consider the alternatives and choose what we think is the best one to focus on based on these predictions. As an example, per our discussion in Sections 7.3.3 and 7.3.5, it should be clear that even our earlier strategies used a type of online optimization to choose which predator/prey to focus on. Moreover, for the strategy in Equation (7.8), we used a type of prediction in determining which was the best predator/prey to focus on (there we predicted which T_i would be highest after a delay, scaled that prediction, and then used it to decide which predator/prey to focus on). For that strategy, the information we used was quite simple, and only incorporated some information about delays in the organism and environment.

If environmental or organism information is available, it can be used to plan what to attend to.

Hence, in a limited way we have already considered the use of planning concepts in attentional strategies. Here, however, we will consider two explicit ways to incorporate more detailed information about the environment. We invite the reader to evaluate the performance of these strategies in Design Problem 7.6.

Attentional Strategies that Use Predator/Prey Behavioral Characteristics

The strategies considered up to this point do not incorporate a significant amount of a priori information that may be available about the likely timing of predator/prey appearances. For example, if the organism has identified the predator/prey *type*, it may have a good guess of when the next appearance time will be, or if it has observed a fixed pattern of appearances in the past, it may have a guess of when it will appear again. Without using such a priori information, the attentional strategies may focus on a predator/prey even though it is unlikely that it will appear or be found for some period of time, and during this time, the organism could more profitably search for and detect other predators/prey.

How can such a priori information be incorporated? We simply provide a few ideas here. First, suppose that we use a "certainty of appearance" function

$$C_i^k(t, t_i^k)$$

for each predator/prey i, which is defined along the time-line $t \geq t_i^k$ starting from the time t_i^k when predator/prey i was last detected (i.e., from the time that the predator/prey was detected for the k^{th} time). Suppose that this function has values in the range of $[0, 1]$, with 0 representing that it is unlikely that there will be an appearance, 0.5 representing uncertainty about whether there will be an appearance, and 1 representing that you are certain that there will be an appearance (based on a priori information). Now, suppose we define a strategy that at each decision point simply picks the predator/prey to focus on that is most likely to appear (and perhaps taking into account any delay in switching focus to a different predator/prey). See Figure 7.15.

In Figure 7.15, notice that there are appearance certainty functions for four predators/prey. Predator/prey 1 is predicted to appear with a higher frequency, and the width of each of the humps quantifies the certainty of occurrence of appearance; hence, appearances are most certain at the peaks. Notice that predators/prey 2 and 3 are predicted to have similar (lower) frequency appearances, but the precise timing of the appearances is not as certain and this is quantified via the spreads of the humps being larger. Predator/prey 4 is predicted to be a lower frequency illuminator, but the certainties of when the appearances will occur is similar to that specified for predator/prey 1. Note that the parameters defining the $C_i^k(t, t_i^k)$ functions (e.g., the points where the peaks occur and the spreads) could be estimated in some situations by some other cognitive subsystem, and then the $C_i^k(t, t_i^k)$ functions used by the attention scheduler could be changed. Finally, note that these certainties could be scaled by predator/prey "priorities" so that the strategy could choose to focus on the highest priority predator/prey that is likely to produce an appearance.

Will this result in a stable strategy? No, not if that is all that is used in the attentional strategy. It could be that you have bad a priori information so that bad guesses are made and the appearances are never found for a predator/prey and so the length of time that it is ignored goes to infinity (representing that

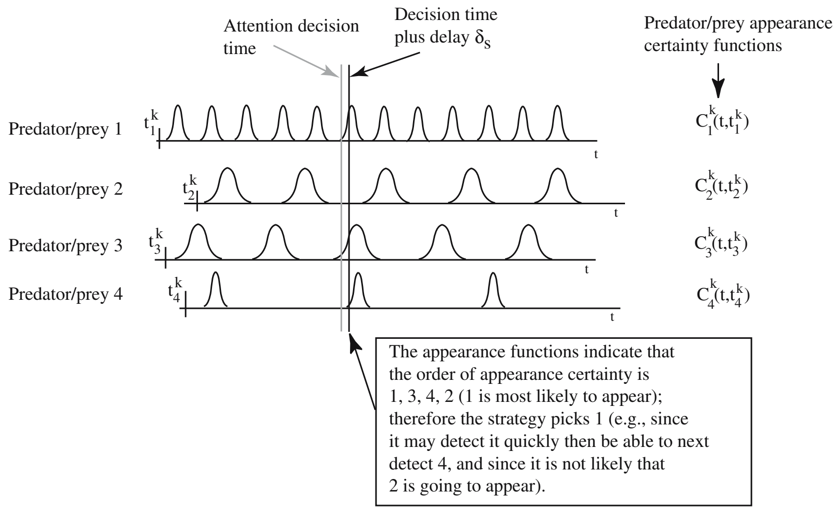

Figure 7.15: Attention strategy that exploits information about likely times of appearance of predators/prey.

ultimately it knows nothing about the predator/prey). We can, however, use the USM of the previous section to ensure stable operation. Moreover, if the $C_i^k(t, t_i^k)$ represent good predictions about the predators and prey, it is possible that very good scheduling performance can be achieved.

Attentional Strategies Based on Model Predictive Control

In most engineering applications we can simulate, to a reasonable degree of accuracy, the domain in which we make decisions. For instance, in this chapter we have simulated the predator/prey environment in Section 7.4. The actual predator/prey environment is certainly somewhat different from what our simulations would lead us to believe. Let us suppose, however, that we can simulate the predator/prey environment reasonably well, at least in its broad characteristics. Furthermore, suppose that the organism can simulate this model of the predator/prey environment *in real-time* in some cognitive module. Would such a simulation provide useful information to help decide which predator/prey to focus on? Below, we study this question by providing one way to incorporate a simulated predator/prey environment into a scheduling strategy.

Suppose that we use the model of the predator/prey environment to predict how the organism will perform using different strategies or orders of focusing on predators/prey. Suppose that we use the model to predict M different behaviors that result from M different candidate sequences of predators/prey to focus on of length N_h (a specification of a sequence of N_h predators/prey to focus on). The strategy is shown in Figure 7.16.

As shown in the figure, for the MPC strategy we rank order the M different

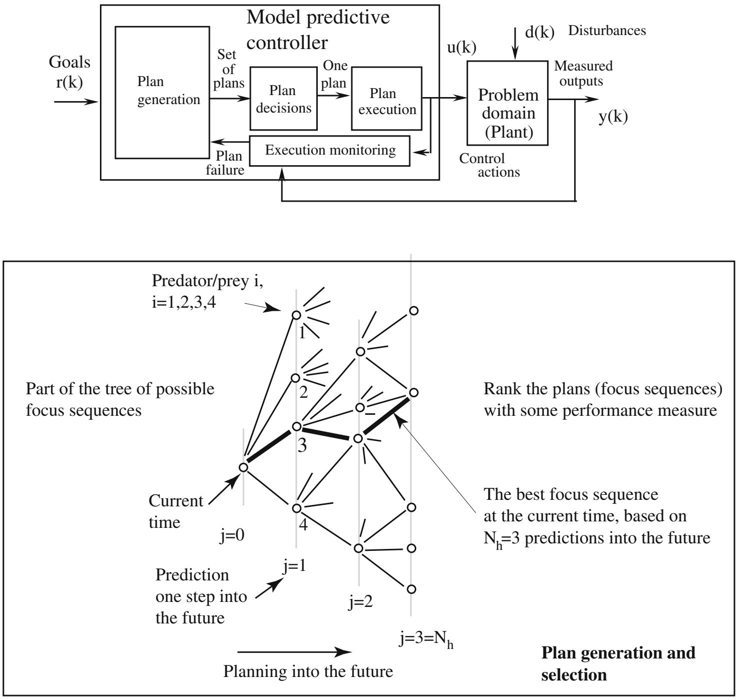

Figure 7.16: Model predictive control (MPC) for use in an attentional strategy.

predator/prey focus sequences and choose the best one, and then we focus on the predator/prey specified as the first one to focus on in the best predator/prey sequence. The process repeats at the next decision point, i.e., when that predator/prey is detected. You can think of the MPC attentional strategy as a more sophisticated version of the attentional strategy discussed in the last section. We think of N_h as specifying a "receding horizon" or length of time we predict ahead in time. For a very uncertain predator/prey environment it typically does not make sense to make N_h very large since the predictions will typically become more inaccurate as we predict farther ahead in time. If, however, your model is good and you have sufficient computational resources you may want to predict into the future for longer periods of time so that the best possible predator/prey is chosen to focus on.

Clearly, if information was gathered online, you may be able to profitably update the model that is used in the MPC strategy (this would then result in the incorporation of learning and planning into attention). Moreover, it is not difficult to incorporate a predator/prey priority scheme. Will MPC-type strategies result in stable scheduling? Probably not. However, once again we can use the USM to ensure that we obtain stable operation.

7.6 Attentional Systems in Control and Automation

In this section we will overview how attentional systems and multisensor integration can be used in control and automation. For more information on each of these topics, see the "For Further Study" section at the end of this part.

7.6.1 Attentional Strategies for Control

In this section, we briefly explain how to augment the control strategies considered so far with attentional mechanisms. Later, in Chapter 9, we briefly discuss relationships between learning and attention and in Section 9.4.5, we discuss how to augment adaptive (learning) controllers with attentional mechanisms.

At the neural level, attentional mechanisms can be implemented by neurons so that an organism focuses on the most important aspects of its environment in achieving a control task (e.g., stimulus-driven attention reorienting that is implemented in a network of neurons). There has been a variety of neural network models introduced for attentional systems, and some of these have been experimentally validated to a certain extent. Some of the models have incorporated the hierarchical aspects of attention, while others have illustrated how attention is integrated with visual processing such as object recognition. Here, we do not investigate neural network models for attention, but in Design Problem 7.5 we provide some references and invite the reader to do so.

Attentional strategies can be employed in rule-based planning and learning controllers.

Typically, the central issue in augmenting a fuzzy or expert controller with an attentional mechanism is to add a mechanism that manages the matching process since that is typically the most complex part of those systems, and the part where sensory data are processed to determine how they should be used. The attentional system in this case could try to prune the number of rules that are on at any one time based on contextual information that is gathered. For instance, suppose that you have a controller with many inputs (e.g., 1000 or more). In this case, you could define priorities for your control objectives and then you could only consider inputs that help you to meet those objectives, or you could process the inputs to capture the essential features. This would be a supervisory strategy that managed the flow of input information so that the computational complexity is reduced. This strategy is shown in Figure 7.17.

To achieve "attentive planning," the ideas for integrating planning and attention in the last section could be useful, or, the attentional system could prune projections into the future (as in Figure 6.2) since that is often the most computationally complex part of the planning process. This is pictured in Figure 7.18. Goals, hard constraints, and other inputs may provide the information for how to prune. Attention can make the complex problem of predicting the many ways that the system can behave in reaction to different sequences of inputs, but it could result in a performance degradation in control performance. Essentially, attention tries to reduce complexity to a manageable level, without sacrificing too much performance.

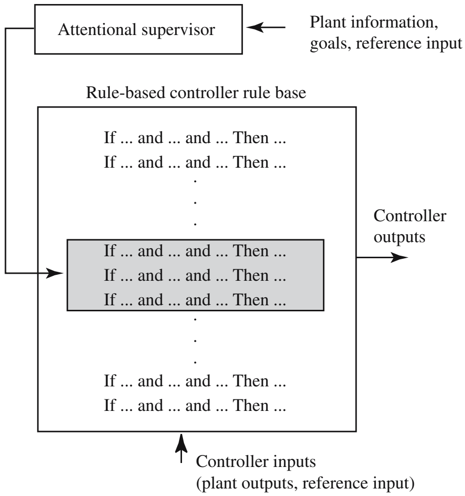

Figure 7.17: Attentional strategy for rule pruning for rule-based control.

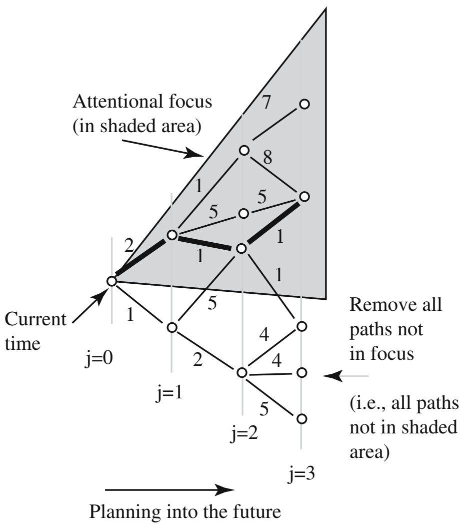

Figure 7.18: Attentional strategy for plan pruning.

7.6.2 Filtering and Focusing: Multisensor Integration

In complex highly automated systems, it is often necessary to use multiple types of sensors for obtaining information about the environment (plant). For instance, a mobile robot may need sensors for velocity, acceleration, yaw, etc. It may also need a vision system for obstacle avoidance, coupled perhaps with radar or an ultrasonic sensor for reliability in achieving obstacle avoidance. The robot must decide how to combine this information for object recognition, decision-making, and other tasks. For some tasks it may ignore some sensor data, and pay attention to other data. For other tasks it may "fuse" data from two or more different sensors. The general task for a "multisensor integration system" is to distill the most useful information from the suite of sensors. A general sensor integration system is shown in Figure 7.19.

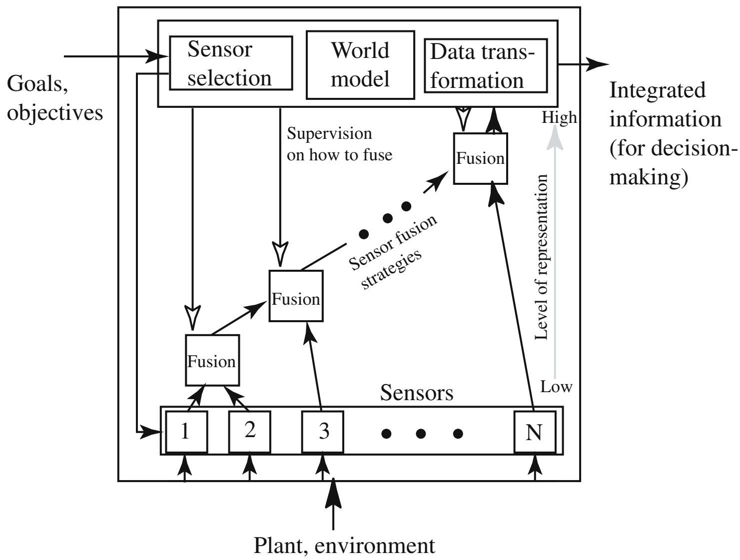

Figure 7.19: Multisensor integration and fusion (adapted from [339], © IEEE, used with permission).

Sensor fusion and integration are closely related to key functionalities in attentional systems.

Here, we see that there are N sensors, each possibly of a different type, or copies of one type of sensor (e.g., for reliability purposes). There is a sensor selector that decides which sensors should be enabled. Then, there is a sensor fusion strategy. In Figure 7.19, we show one strategy where information from sensor 1 is fused with information from sensor 2, and then that fused information is fused with the information from sensor 3, and so on. Other strategies are also possible (e.g., having two fusion strategies combine information from two sensors each, then you could have another fusion approach for the fused information from those). There are a wide variety of methods for multisensor fusion (e.g., Kalman filtering, Bayesian estimation, etc.), world modeling, sensor selection, and data transformation. The interested reader should consult the references in the "For Further Study" section at the end of this part.

The fusion strategies may have guidance from the higher level functionality. The fused information is passed to the higher level and may be stored in a "world model" (a representation of aspects of the environment that are useful for decision-making to reach the goals, but which may also help guide the overall strategy as to how to fill information that is needed). To achieve world modeling and ultimately sensor integration, we will often have to also perform data transformations. Ultimately then, multisensor integration systems do possess some key features of learning, and that topic is covered in Part III. Next, note that there is a general process of filtering that is naturally involved in sensor fusion where some information is discarded and other information is derived by combining data. This results in a "low" level of representation early in the fusion process and a "high" level of representation at the end where the most useful information has been obtained.

There are several possible advantages to integrating information from multiple sensors. For instance, some sensors may provide *redundant* information which can reduce overall uncertainty about what is being sensed, or it can provide for fault tolerance in case a sensor fails. Sometimes information is *complementary* in the sense that it may allow, via appropriate processing, for the perception of some objects that could not be perceived otherwise. Sometimes, multisensor integration can speed up the overall process of decision-making by providing the proper information faster. Other times, it may be possible that using multisensor integration strategies will result in a less-expensive system.

Finally, we note that the focus in this chapter is largely *not* on attentional strategies for control or on multisensor integration, but on how to use control concepts (scheduling for resource allocation) for attentional strategies.

7.7 Exercises and Design Problems

Exercise 7.1 (Simulation of Attentional Strategies):

(a) Simulate all the attentional strategies in Section 7.4, reproducing the results found there.

(b) Let $N = 3$ and use

$$\delta^1 = 0.9,\ \delta^2 = 1,\ \delta^3 = 1.2$$

Synthesize sequences similar to those shown in Figure 7.7 that satisfy these constraints (make the appearances periodic). Choose $\delta_s = 0.03$ and

$$a_1 = 0.3,\ a_2 = 0.2,\ a_3 = 0.1$$

Simulate the three attentional strategies studied in Sections 7.4.2 and 7.4.3 and evaluate the performance of each attentional strategy. Tune the w_i parameters to obtain as good performance as you can via manually tuning these parameters.

Exercise 7.2 (Stability Analysis of Priority-Based Attentional Strategies): Prove that the policy defined by embedding priorities via the approach in Section 7.3.4 into the strategy defined by Equation (7.1) is stable if appropriate conditions are met (state these, and show each step in your proof). Provide explicit ultimate bounds on the T_i values.

Exercise 7.3 (Stability Analysis of an Attentional Strategy): Prove that the attentional strategy defined in Equation (7.8) is stable if appropriate conditions are satisfied (specify the conditions and show each step in your proof). Provide explicit ultimate bounds on the T_i values.

Design Problem 7.1 (Tuning Attentional Strategies with Priorities): This problem focuses on how to tune attentional strategies to improve their performance.

(a) Specify priority parameters $p_i \in P$ and explain how to embed a priority scheme into the attentional strategy in Section 7.3.2 using the ideas in Section 7.3.4. For a specific set of priority parameters, develop a simulation of the priority attentional strategy and evaluate its performance for the scheduling problem defined in Section 7.4. Tune the priority parameters to try to improve performance, where you measure performance by the time average of the average values of the lengths of times waited.

(b) Next simulate the strategy given by Equation (7.8) and tune the w_i parameters to obtain better performance, as measured by the time average of the average values of the lengths of times waited, than what we obtained in the chapter. Compare the performance that you obtained to that which you obtained in (a).

Design Problem 7.2 (Stable Attentional Strategy Design):

(a) Suppose that you consider the set P of labels for the predators/prey as specifying the sequence that they should be focused on (and suppose that this sequence is fixed a priori by the labeling). Suppose that you define a policy that at each decision point simply picks predator/prey $1, 2, \ldots, N$ in sequence, and after it finishes with predator/prey N, it returns to predator/prey 1 and repeats the process. Will this result in a stable attentional strategy? Why? Why not? Can you generate a counterexample to stability, or provide a proof of stability that does not use the USM?

(b) Can you define an attentional strategy that will result in stable operation, but is different from the others discussed in this chapter and does not use the USM? Specify the strategy and prove stability.

Design Problem 7.3 (Design of Universal Stabilizing Mechanisms): For the scheduling problem in Section 7.4, employ the attentional strategy defined in Design Problem 7.2(a). Augment the strategy with the USM. Simulate the strategy for various choices of USM parameters and explain the effects of these parameters on attentional strategy behavior and performance (measure performance by the time average of the average values of the lengths of times waited). Be sure to simulate the attentional strategy for a sufficient period of time so that the performance measures represent the long-term performance of the attentional strategy. To get accurate performance measures, do you need to repeat the simulation many times with different sequences of choices of predators/prey to focus on?

Design Problem 7.4 (Design of "Optimal" Attentional Strategies): This problem builds on Design Problem 7.1 by exploring systematic ways to pick the best attentional strategy parameters. Choose a *stable* attentional strategy (you may use the USM) that seems to have the potential to obtain a better value for the time average of the average values of the lengths of times waited than the one in Section 7.4.2. One approach to this is to tune the priority or weighting parameters for the attentional strategies in Design Problem 7.1 to try to obtain better performance. Tune the parameters of the attentional strategy with a goal of obtaining a better value for the time average of the average values of the lengths of times waited than the one in Section 7.4.2. Hint: You may want to produce a systematic approach to tuning the parameters of the scheduler rather than just manually tuning them. One approach to do this is to use ideas from the "response surface methodology" discussed in Chapter 15. A simple version of this approach is to simply create a grid of attentional strategy parameters and simulate the strategy for each point on the grid (which can take significant computational resources) and pick the parameters that correspond to the best performance. Another approach would be to use the "simultaneous perturbation stochastic approximation" algorithm that is studied in Chapter 15.

Design Problem 7.5 (Neural Models of Attentional Systems)*: There is research in the literature on how to develop neural network models of attentional mechanisms and this problem studies the simulation of attentional systems via such models.

(a) For background reading, read the article [372]. Search the literature on this topic to supplement this study.

(b) Implement code necessary to study the attentional system and reproduce the simulations shown in [372]. Focus on the simulation of the "spotlight" view of attention.

(c) Explain how such an attentional mechanism may be useful in a control system. Identify at least two ways in which it can be used.

Design Problem 7.6 (Attentional Strategies Based on Planning and Learning)*: In Section 7.5.3 we introduced two ways to use planning concepts in attentional strategies. Here, you will completely specify such a strategy, simulate it, and evaluate its performance.

(a) Using the ideas in Section 7.5.3, develop an attentional strategy that incorporates planning concepts. You do not have to precisely follow the methodologies specified earlier; you can invent your own method. Specify the attentional strategy, explain what environmental/organism information it needs in order to predict how the attentional strategy will operate, explain what cost function will be used to select a single sequence of predator/prey focuses (plan) from the set that is generated, and explain how the overall approach seeks to improve attentional performance. Specify the strategy in a way that will ensure that it is stable (you may use the USM).

(b) Specify a single performance measure that you would like to optimize. Develop a simulation of the attentional strategy that you specify in (a) and tune the strategy to try to optimize your chosen performance measure. You should use a scheduling problem and performance measure similar to the ones in Section 7.4.

(c) Next, expand on your strategy by incorporating a method to learn the model that is used by your planning strategy to predict. Repeat (b) for this strategy.

Design Problem 7.7 (Cooperative Attentional Systems)*: Suppose that there are M agents, each with an attentional system given by the model used in the chapter. Suppose that they are seeking predators and prey, but that they do so cooperatively in the sense that they identify N predators or prey and then cooperate on paying attention to them. With cooperation we expect that there will be an increased "capacity" to pay attention.

(a) Define two cooperative attention strategies. For instance, suppose that the $M < N$ agents act autonomously but share an "unattended" set of things that are not paid attention to at the current time. There is then a corresponding set of predators/prey that the group of M agents is attending to. A decision strategy can be defined in terms of what each agent does at its decision times. For instance, it may "check out" (using a mutual exclusion strategy) the unattended set and pick a particular predator/prey to focus on; then it can "return" the new unattended set to the others. It can then focus on that predator/prey until it is detected. The agents would then make all their decisions asynchronously. What predator/prey should be chosen from the unattended set? Mathematically define two strategies for the agents to make these choices.

(b) Simulate the cooperative attentional strategy and show plots as we did in the chapter to evaluate their performance (e.g., relative to the $M = 1$ case).

(c) Augment your strategies with the learning/planning methods you studied in Design Problem 7.6 and then evaluate their performance in simulation.

(d) Find conditions under which the strategies of (a) will result in stability in the sense that it was studied in the chapter.