

第三章 設計代理人的決策系統

決策過程將會經過模型中的記憶、情緒、學習模組來處理，學習模組同時也記錄著模組之間的活動。記憶機制包含三種模組：長期記憶(long-term memory, LTM)、短期記憶(short-term memory, STM)、以及工作記憶(working memory, WM)。每個記憶模組對於訊息屬性負責不同的衰退值(decaying)快慢以及優先權(priority)的高低，並將處理過的訊息依照程序傳值到下一個目的地：其中長期記憶與短期記憶模組屬於軟體前期的資料程序，而工作記憶模組是屬於實體的(physical)電腦硬體資料流程。文中將描述記憶及情緒模組的概觀，如何個別地呈現及相互合作，更詳細的研究論述與計算模型可以在下列(Sandra, 2003) (Lee, 1999) (Ortony, Clore & Collins, 1988)等研究中獲取。在一開始的資料流程，資料將第一個被資料分析模組(information analysis model)所處理，它將根據記錄在長期記憶區(LTM)中的資料來區分每個資料的種類以及優先權高低，同時長期記憶區(LTM)中的資料也經常性地被周遭模組來修改內容。因此，在接下來的兩小節中，我們討論情緒模組如何運作以及記憶機制如何來影響決策的過程。

3.1 決策系統中的短期記憶



短期記憶(STM)的功能就像資訊有限的容量以及生命週期(Lerman & Galstyan, 2003) (Peters & Sullivan, 2002)。短期記憶模組位於長期記憶(LTM)模組與工作記憶(WM)模組之間(圖示 3-3 中 path 8, 9)，短期記憶就像是兩模組之間的通道，它同時也從學習和資料分析模組(information analysis model)接收資訊。短期記憶模組中包含了對於事件的計數器，當計數器的值到達特定標準時，短期記憶就會把該事件資料經過學習模組轉送至長期記憶模組(圖示 3-3 中 path 6 to 7)。而通往工作記憶模組的資料則是已經經過所有模組評估後準備進入實際執行的工作佇列(queue, FIFO)，資料在工作記憶中將被電腦所執行 (Liew & Gero, 2002)。

儲存在短期記憶的資料將會被評估其存取的優先度，藉由每項記憶資料的存取次數而定。經由放置在短期記憶模組中獨立的計時器(timer)，每個資料都會附加一個衰退值(decaying value)，衰退值將會逐漸地減少最終使得該項資料消失，除非資料被其他模組所存取使衰退值恢復預設值。如果該項資料被存取的次數高於設定的門檻值，資料將送至長期記憶放置以提供進一步的存取(Liew & Gero, 2002)。存放在短期記憶的資料也被用於決策的過程，比較對照這些已經定義清楚的值並做出決定。以下是整理過數個相關於短期記憶的特質：

- 記憶功能的運作是動態的存取資料而不是將資料靜態地存放於記憶空間
- 當決策過程所需的項目不存在時，記憶功能將提取相關項目的平均值
- 短期記憶資料的傳送規則不是預先定義好的，系統開始運作時資料屬性開始累積

3.2 決策系統中的長期記憶

放置在長期記憶(LTM)模組中的資料比在短期記憶模組中保存時間更加持久(Peters & Sullivan, 2002)，其功能相似於資料庫(database)的性質。資料在長期記憶中一樣具有衰退值，但衰退值遞減的速度非常緩慢。長期記憶模組同樣也具有存取檔案優先高低(accessing priority)的分類，使得檔案存取時的效率能得到最佳化。當系統開始運作的時候，在長期記憶中的資料就會被取用作為比對新進資料用途，作為排除錯誤或不明確的檔案。另一條從短期記憶至學習機制導入的路徑(圖示 3-3 中 path 7)，將增加長期記憶本身所存放的資料量。在整體模型的功能上，長期記憶提供了增進模組之間累進的互動基礎，也提供代理人在行為上的依據。

這裡有一些功能是與長期記憶模組合作後才能進下一步的動作。由於長期記憶模組的運作方式類似於資料庫，其中學習、情緒、資料分析模組都需要以長期記憶為基礎的運作方式。當資料剛送至長期記憶模組時，資料將根據他們的屬性與目的進行分類，圖示 3-1 呈現了分類的流程。資料經過分類後將送至不同的地方儲存，而每一個傳送過程都有計數器在累積次數，其中有一些類目被標示為特定用途所以有自己的計數器。這些計數器功能會與模組本身的存取優先排列相互合作，和模組以外的學習機制也有相同的合作關係。

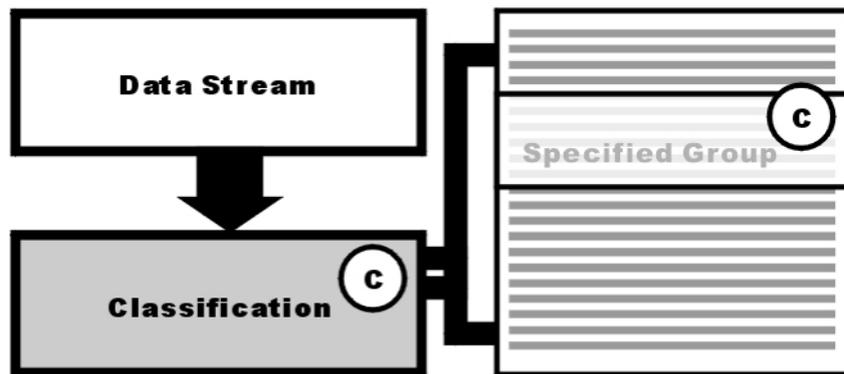


圖 3-1: 資料分類流程圖解

經過資料分類之後，存取優先權排列將是下一個資料處理程序。根據被其他模組存取的頻率，資料被安排不同的優先等級以提升存取速度。經過一連串的操作過程後，長期記憶模組會將累積的計數值傳送至學習機制中(圖示 3-3 中 path 7)，其中包含每個事件的使用率以及事件與事件的聯結關係。學習機制取得數據之後，往後當代理人遭遇類似情況時，學習機制將回傳已知數據供長期記憶模組預先載入可能需要使用的函式，以節省搜尋存取時間。雖然精確有效率的優先權排列方式仍然有許多討論空間，簡易的評估過程已經能幫助目前的需求。以下是幾個關於長期記憶的重要特質：

- 存放在長期記憶中的資料仍然具有衰退特性，而不是永遠保存不滅
- 根據使用的頻率，資料在長期記憶中擁有存取權的高低
- 長期記憶中的搜尋存取規則是影響系統效能其中一項重要因素

3.3 決策系統中的情緒模組

在我們的觀點來說，情緒功能提供了決策系統一個錯容以及彈性的方法解決問題，並在模組之間的合作上提供得到更多經驗的機會(Kumar & Cohen, 2000)。隨著虛擬環境複雜度的提高，代理人必須藉著自己的決策過程來面對許多事件發生，而且事先不可能針對所有的環境細節作分辨；當代理人不能完成或到達預定的事物，無法預期的錯誤可能會發生。除了事先定義錯誤的解決方法而避免錯誤發生之外，提供一個錯容性(fault-tolerated)的方法來避免這樣的處境是被需要的，而情緒模組則給予了上述問題一個良好程序。舉例來說，當未知的錯誤發生時，情緒模組將錯誤的訊息進行分類並儲存成一個臨時數值，藉由數值再去尋找可能的解決方法，系統將忽略這個錯誤繼續正常運行，接下來我們將更詳細的討論。

當資料分析模組(information analysis model)完成資料處理之後，緊接著將資料傳送至情緒模組來進行下一個程序(圖示 3-3 中 path 3)。情緒模組將評估新進資料並檢查其他函式(function)的覺醒值(arousal intensity)，是否有必要呼叫其他功能來加入評估，然後將所得的經驗傳送至學習模組(圖示 3-3 中 path 5)。

3.4 情緒模組的評估

情緒模組內部的評估方法主要分成兩個部份，一個是處理定義明確的資料，另一個則是處理不明確的資料(請參考圖示 3-2)。新進資料將依據分析函式來分辨檔案明確與否，而分析函式依照是否能在長期記憶(LTM)中找到該項符合的數值來判別。如果資料屬性為定義明確，該資料將直接傳送至下一個目的地。普遍來說，清楚的資訊對於情緒模組來說是正面的，而不清楚的資訊是負面的，負面的情緒狀態將產生類似憤怒、恐懼、緊張等的行為表達，並使得覺醒值(arousal intensity)提高而開啟其他模組功能。在情緒相關的評價方法中，我們採用的是二分法來簡化整體模組程序的複雜度。

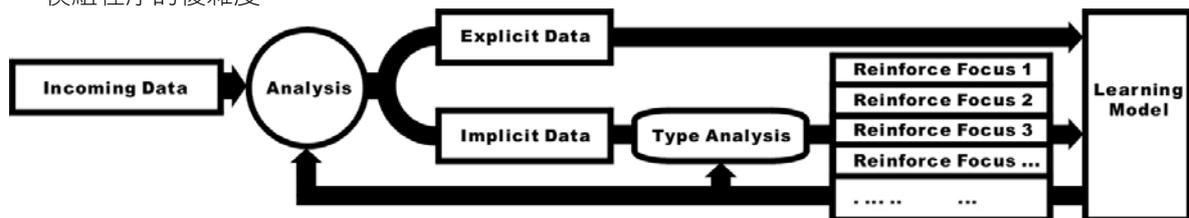


圖 3-2: 情緒評估資料過程圖解

當資料被認定為不明確的屬性時將會被送至另一條管線作處理，該項資料將被決定哪一個連結模型與它有關。連結模型代表該項資料有可能涉及的範圍，列舉出可能找出解答的相關函式，嘗試從有限的條件中得到最佳解。舉例來說，假設代理人活動需要能量作為供應，當能量快用完的時候它會嘗試關閉一些比較不重要的系統功能並集中在搜尋功能以求生存機會。當模組在有限的條件下依然無法找到合適的方法時，代理人將停止對於該項資料的回應並將模組的工作紀錄送至學習機制中，以作為未來影響模組判別的參考經驗。

3.5 決策系統中的學習模組

使代理人在虛擬環境中具有適應性的行為，學習機制(learning mechanism, LM)是我們面對的另一項重要議題。在本論文中我們將不會談論關於學習方法的進階應用，而是一些簡單初級的自我學習。站在進化的觀點上，學習是唯一的方法使代理人在決策能力與行為功能上有實質上的增進(Kozierok & Maes, 1993)。我們使用簡單的統計學與模糊邏輯作為學習機制中主要的兩個功能，下面是幾個學習機制的重要特質：

- 基於系統所記錄的歷史(STM and LTM)，學習機制預測使用者的意圖或行為
- 所有在學習機制中所設定的參數在系統一開始的時候都是中立的
- 根據預測數值的多少，預測標準是動態建立的

學習機制將包含回饋學習與監督學習的部份特質，能夠經由嘗試與錯誤中學習新的經驗，以及經由統計的辨識方法來熟悉週遭環境的變數。在大部分的情況下，我們希望所有的資料都能夠經過學習機制統計使得經驗數值更加客觀，基本的概念是新進資料與已記錄的歷史經驗相互比對，以求得適切的系統反應。圖示 3-3 中有三條路徑與學習機制相連結(path 5, 6, 7)，各自收集有效的工作紀錄來強化學習功能，而這些模組也被學習機制所影響。在圖示 3-3 所提到情緒模組與學習機制的關係(path 5)，學習機制將紀錄從情緒模組管線輸出的特定資料屬性及其事件結果，同時學習機制也會更新情緒模組的判斷標準，包括不明確資料管線的搜尋條件(請參考圖示 3-1)。在一些特殊的狀態下，代理人必須在不明確的決策過程之中做出決定，此時學習機制將會評估可能的事件範圍，計算出數值的平均範圍而產生可能的目標；如果條件差距過大，代理人對於該項事件將暫不回應，並維持先前的目標或設定。學習機制收集所有事件或決策過程的計數值，當數值超過門檻時便把該項資料存入長期記憶(LTM)中，作為資料比對的依據。

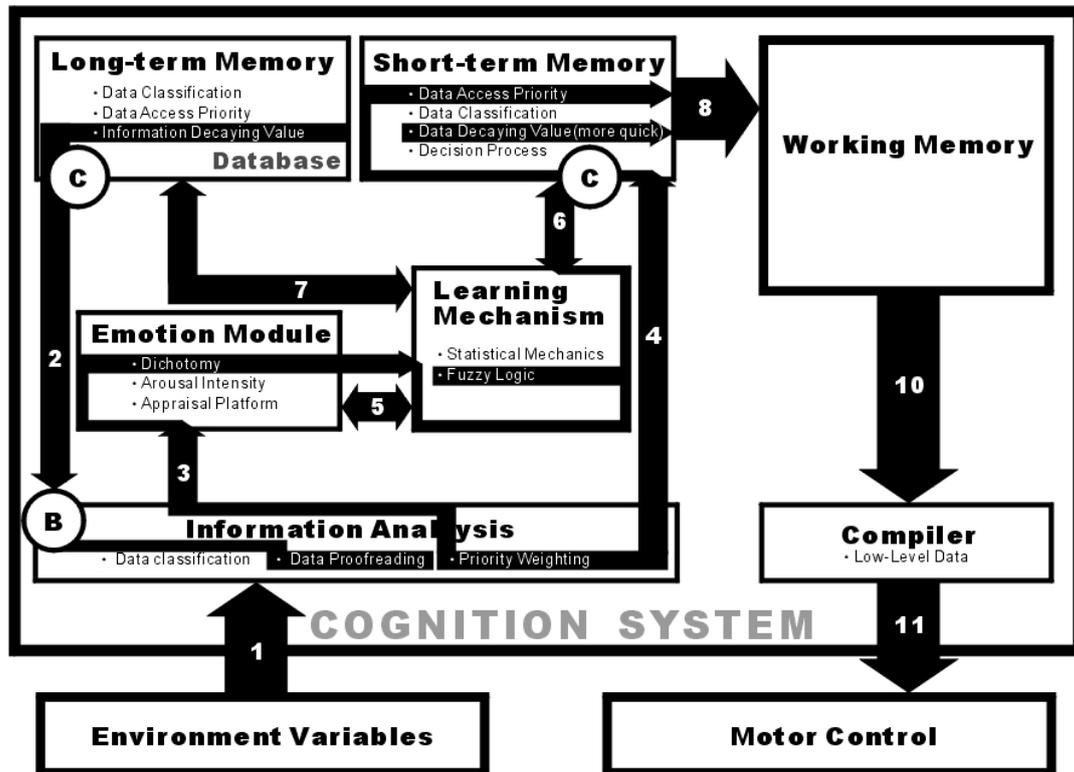


圖 3-3: 認知系統流程圖解

3.6 模組之間如何相互合作

本研究所提出綜合性模型將與實作部份的模擬程式相互對照與驗證，提出模組之間相互合作的實際參考；模組之間不僅各司其職，同樣也是相互影響的關係。模擬程式將依照圖 3-1 所提出的路徑與決策分支作為程式撰寫的參考依據，其中並包含代理人面對環境時的各項變數與參數設定。以下將舉三個情境範例作為決策互動過程的說明。

代理人所處的環境包含許多訊息，其中包括代理人身處站立的位置；此區塊包含了相關的地理訊息，包括地理座標的位置、地形坡度、所劃分的用途屬性...等等。假設代理人所停留的位置在人行道，一開始他會將人行道這個地理訊息送至環境分析模組進行環境變數分類，標籤每種地理資訊屬性；然後開始進行與長期記憶模組(資料庫)作資料比對，檢查是否有已經存在的相同訊息；假設比對的結果正確，由於此項資料對於代理人本身並沒有相斥的成分，因此資料將經由正常程序傳送至情緒模組。在情緒模組中資料將經由二分法來判別是否須要進一步的模組評估(請參考 3-4 節)，並檢查模組中數值的門檻值，假設此項訊息無誤，便將訊息傳送至下個目的地：學習模組。學習模組將此項訊息列入分類的統計數值中，然後將訊息傳送至短期記憶模組。訊息一傳送至短期記憶模組便將開始計算他的停留時間，並開始比對之前是否有相同的訊息出現過，以作為存取的優先考量；訊息在短期記憶模組中留下資料屬性之後便送至執行單元中運作，呼叫適當行為函式作為環境的對應；人行道此項地理資訊經過代理人的決策過程後將產生行為，決策過程告知代理人所處的這塊區域可以行走或奔跑(請參考圖 3-4)。

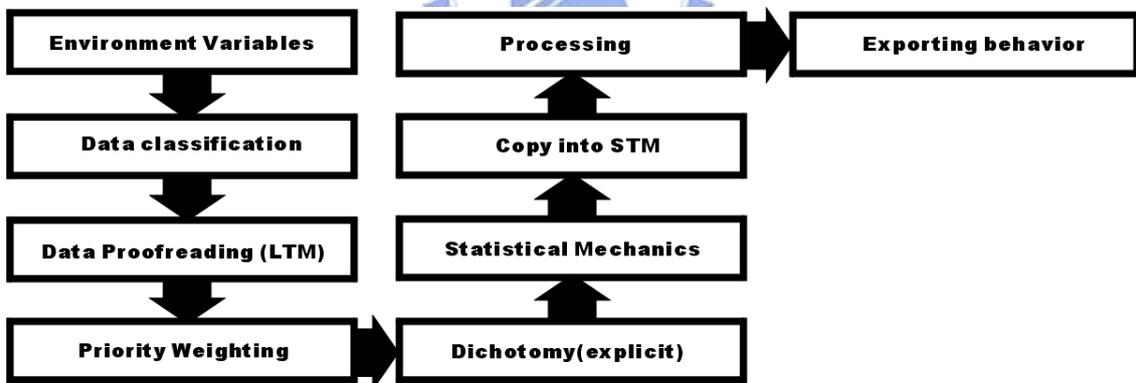


圖 3-4: 第一個決策過程範例 - 人行道

第二個範例將假設代理人遭遇已知危險狀況時所採用的決策過程。當代理人所身處的範圍是已知的危險區域時，地理資訊將同樣送至資訊分析模組進行處理。經過資料的分類之後訊息傳送至資料比對函式，比對之後從長期記憶模組中得到此項訊息是危險的屬性，便將訊息傳送至情緒模組相關的優先權評估函式；函式確認訊息屬性屬於危險之後，便直接將訊息傳送至短期記憶模組登記該項記憶數值，然後經由經驗法則為依歸(rule based)的行為模式來避開或減低代理人受損的程度(請參考圖 3-5)。由於在危險的狀態下代理人必須緊急的針對問題做出回應，在有限的時間內以減輕環境所造成的傷害，代理人將直接省略決策過程中一些模組執行程序，並集中處理解決的方法。此項事件的決策過程突顯了情緒模組所擁有的彈性特質，尤其在龐大系統架構下所呈現的效果更加明顯，但相對的早期經驗法則規劃就更顯得重要。

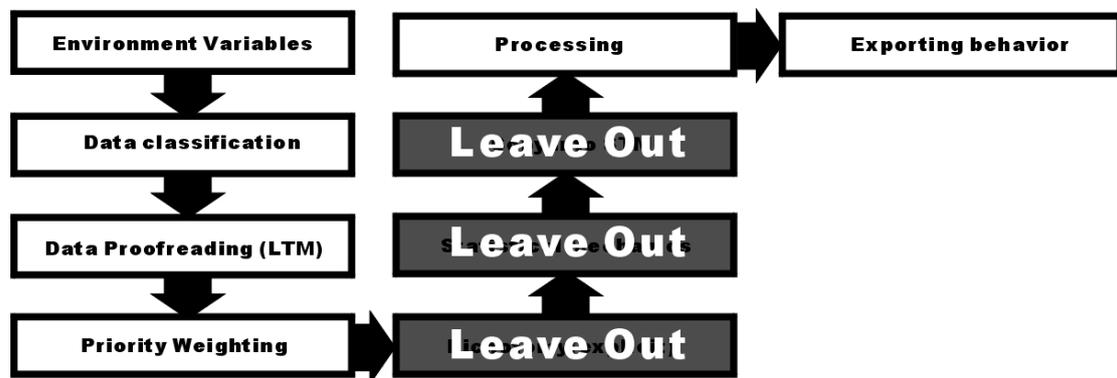


圖 3-5: 第二個決策過程範例 - 危險變數



第三個範例假設代理人遭遇不明確事件的決策過程，如何透過模組得到學習的經驗，並在未來利用調適性行為模式來避免相同事件的發生。同樣一開始資訊將送至資訊分析模組進行處理，由於是不明確的資料型態，將跳過初步的資料分類進入資料比對函式。資料比對函式在長期記憶模組中未找到該項符合的數值，因此將分配至情緒模組二分法中的不明確檔案類型管線進行分析(請參考 3.4 節)；假設函式在搜尋模式中依然無法確定該項資料的類型與範圍，並且沒有到達任何函式的驅動門檻，此時情緒模組將對於該項資料留下困惑(未解決)的屬性值然後送至學習模組。在學習模組中將此項訊息加入未明確的統計數值，並將資料送至短期記憶模組。因為該項資料屬性屬於不明確的類型，所以無法在短期記憶模組中進行資料儲存或其他程序，直接將資料送至執行單元運作，呼叫適當的行為函式；由於在決策過程中該項資料的屬性無法確定，因此經由決策過程所產生的行為將與開始時的狀態繼續保持一致(請參考圖 3-6)。舉例來說，當代理人在路上行走時路線中遇到障礙物，假設障礙物這項物件資料代理人無法辨識，因此經過決策過程後代理人同樣還是繼續維持行走的行為模式，直到與障礙物面對面接觸。

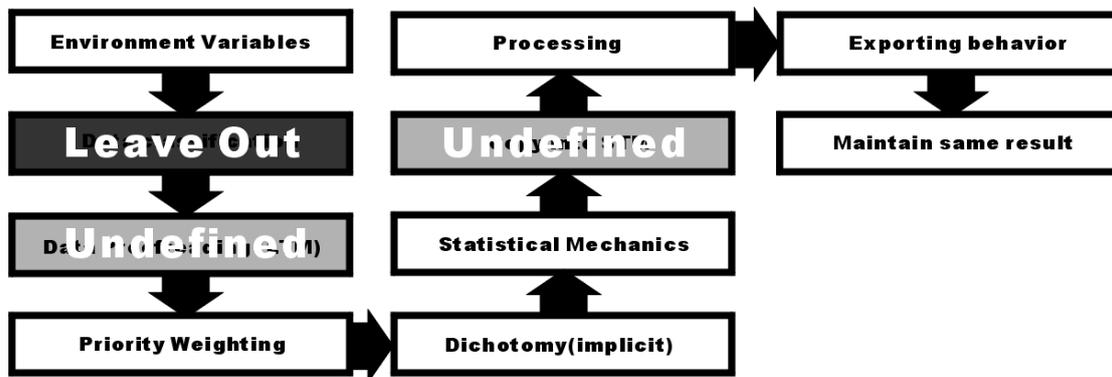


圖 3-6: 第三個決策過程範例 - 未知的變數

代理人所處的虛擬環境將架構在一些既定的規則上，這些規則將類似於現實世界的物理特質，比如像是物件與物件之間的碰撞偵測。物件在虛擬環境中都有著自己的影響範圍與屬性，在初始的狀態下物件之間不會有所交集，除非有環境因素的變動。代理人與障礙物在虛擬環境中同屬於物件的性質，彼此有著不同的功能導向。當代理人與障礙物的影響範圍在虛擬環境中互相接觸時，此時存在於環境中的碰撞偵測函式就會開始運作，計算物件之間彼此的屬性數值與既定規則，然後做出回應。與障礙物的碰撞原則上將使得代理人改變現存的行為模式與運動方向，並且對於代理人所擁有的內定值產生數值上的變化，在代理人的立場來說此項行為結果是屬於負面的。當代理人接收到負面的訊息回饋時，將針對於不久前所做的決策結果重新進行評估。

之前在決策過程中所記錄的不明確資料在此得到解答，決策流程將對於資料重新定義。負面的回饋訊息將直接跳過資料分析模組前往情緒模組與之前所留下的困惑屬性值進行關係連結，情緒模組會依照回饋結果的影響程度對於該項訊息加深其重要性，並且把連結的相關資訊傳送至學習模組來加入明確的統計數值，然後將訊息傳送至短期記憶模組；依照情緒模組所歸類的訊息重要性，學習模組可能直接將訊息結果複製傳送至長期記憶模組中作為前期資料比對的依據，或者是經由短期記憶模組統計評估之後再經由學習模組送至長期記憶模組的另一條路線。不同於之前不明確的訊息類型，短期記憶模組現在可以依照正常程序處理來處理該項訊息，儲存檔案屬性並傳送至執行單元(請參考圖 3-7)。負面的回饋訊息數值將造成代理人行為模式上的改變，作為目前環境狀態的直接反應，並影響後續的決策過程與結果。舉例來說，當代理人與障礙物進行碰撞後產生跌倒的行為模式，此時代理人就能夠辨識出此項原先未知的物件屬性，並且了解障礙物對於本身將在何種情況下造成特定的影響，以至於未來如果有機會再一次面對問題時，能夠提前預防並採取因應的行為對策。

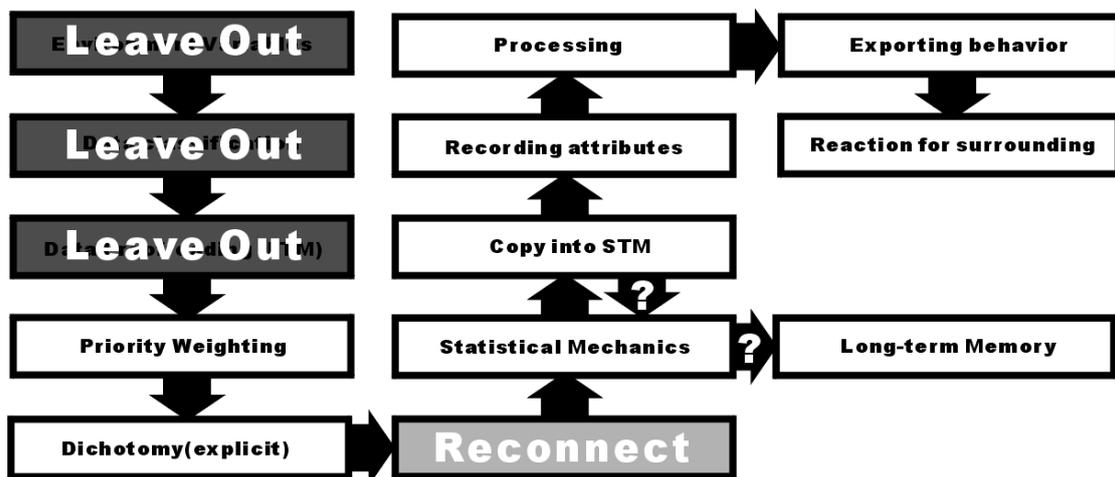


圖 3-7: 第四個決策過程範例 - 重新計算變數

3.7 模組合作的結果

模組之間的合作將使得代理人在決策過程中產生具有彈性並詳細的結果，利用彈性的決策路線讓代理人能有更細緻的行為表現空間。模組之間不僅只是將環境的變數互相傳遞處理，彼此也肩負著系統監督制衡的功效；單一模組本身所能夠涵蓋的功能有限，當目標問題大過於少數模組處理範圍時，整體合作分擔是對於問題最妥善有效的辦法。模組合作也牽涉到彼此的工作監督過程，當單一或局部模組對於問題沒有正確解決的時候，週遭的模組在傳遞訊息時可能就會發現問題依然存在並告知負責單位，嘗試從過程的問題中尋找答案。

代理人於外在行為表現上是模組合作的具體結果，由 3.6 節所提出的假設範例更能了解模組互動過程。代理人在虛擬環境中產生適應性的行為模式來對應於環境變化，並主動偵測所處位置的環境屬性，以避免代理人在虛擬環境中條件規則的牴觸，模組架構的關係請參考圖 3-8。

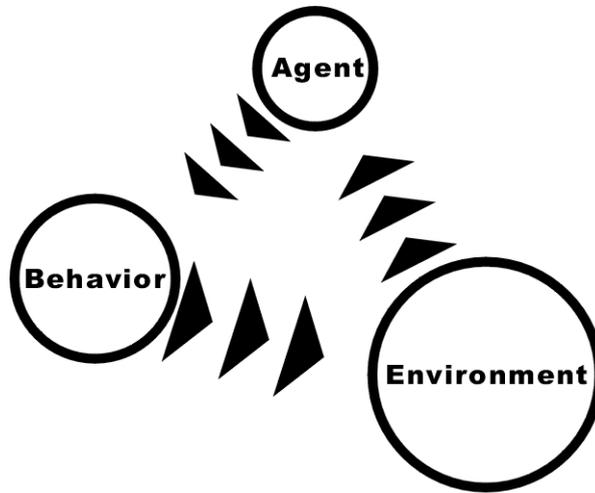


圖 3-8: 模組之間的合作關係圖

