When Being Reactive Just Won't Do

Sharon Wood
School of Cognitive & Computing Sciences
University of Sussex,
Brighton, BN1 9QH, U.K.
email: sharonw@cogs.susx.ac.uk

Abstract

Reactive systems are, generally, highly successful for dynamic uncertain domains. Analysing why this should be so indicates the main criterion for success is that information apprehended about a given situation can render it *certain enough* to reliably inform action. A case is made for informing action by dynamic world modelling for domains where this does not hold. When the issue of attending selectively to situational information is introduced, the case for dynamic world modelling is enhanced. A system incorporating selective perception which dynamically models its domain is described.

Introduction

We are familiar with the idea that dynamic domains pose problems for deliberative planning. The reasons are well-documented (*cf* Georgeff, 1987); to briefly summarise, deliberative planning is generally unable to deliver a timely response for domains where events *outpace* the planning process; they are generally unable to respond to events occurring *during* the planning process; and they are generally unable to handle the *urgency* with which problems must be addressed. To date, the solutions proposed to overcome these problems have generally resided in reactive systems. The main advantages these systems offer is they *can* deliver a timely response to events; they have no planning process to interrupt so they are able both to tackle each event as it arises and to do so with the requisite degree of urgency. Where some element of deliberative planning is preferred, we have seen the emergence of hybrid systems. Often these are multi-layered systems comprising a reactive level and a deliberative level, sometimes more, with the capability to select between the solutions available from each level in order to deliver a timely solution to the current situation (*cf* Ferguson, 1992; Georgeff & Lansky, 1987; Ogasawara, 1993; Sanborn & Hendler, 1988; 1990). Wherever a real-time responsive capability is required, reactive systems, whether incorporated within a hybrid system or not, have generally appeared to offer a solution. If we analyse why this might be the case, I believe there are some interesting insights to be gained.

The Nature of Uncertainty

The analysis I wish to make involves, initially, moving away from a focus on the need to be reactive or fast. However, ultimately, I shall return to this issue as it is pertinent to the analysis. Instead the analysis rests on the relationship of reactivity to uncertainty in the execution environment.

The main problem posed by a deliberative approach to planning in a dynamic domain is that planning takes time. However, time is only a problem in so far as we are uncertain about key features of the execution environment which might inform our planning. In a perfectly known universe, plans may be laid in good time and executed at the apposite moment. This may be stating the obvious, but stating it allows us to make some important observations. The first is that until information about a given situation is apprehended, it is not possible to specify an appropriate course of action (or reaction). The second is that once information about a given situation *is* apprehended, it is certain *enough* to inform action. This second observation is crucial to the analysis of why reactive systems are successful for certain classes of problems or domains.

(Re)action selection depends upon identifying characteristics of a given situation that indicate that a given response is appropriate. This is usually 'hard-coded' as a rule whose antecedent describes situational conditions and whose consequent the (re)action(s) to take. However, a response will be appropriate *only* if it interacts successfully with the situation as it unfolds. In order for this to be the case, the situational conditions which trigger a rule must, to some extent, predict events which will ensue in that situation. There are two ways in which an action may be unsuccessful. The first is through a failure of execution. For example, you turn the handle of a door and pull but nothing happens; or it opens then immediately swings shut again. Herein lies the power of reactive systems: they simply take the current outcome as a new set of conditions which will trigger the appropriate response[1]. The second is that the response made can be described as the wrong response[2]. That is, some other response *should* have been made. It is a moot point whether this is, in fact, any

[1]Herein also lies the argument that *only* a reactive approach will suffice for this kind of problem.

[2]Of course, there are *degrees* of 'wrongness', so the response might be just sub-optimal (compared to, say, a more deliberated response). However, this is not the issue addressed here.

different from the kind of failure described above; however, I think it helps to consider that it is. This is because, with the best will in the world, failures of execution will always occur and their consequences borne. Whereas failure to produce a correct response indicates a failure in the response selection process. This in turn points to a failure to correctly identify or evaluate the situational conditions which led to its invocation; an evaluation which depends upon correctly anticipating situational outcomes on the basis of current observations.

The implication is that reactive systems can be successful, other than by chance, for uncertain, dynamic domains *only* if the information apprehended about a given situation can render predictions about the outcome to that situation *certain enough* to reliably inform action. Without the ability to predict, and hence resolve the uncertainty concerning the outcome to a situation as it unfolds, on the basis of current observations, action selection cannot be reliably performed.

The extent to which this need concern us will depend upon the characteristics of the domain with which the system must interact. As pointed out earlier, if the choice of action is incorrect and the system unsuccessful then the reactive approach can overcome this by evaluating the new situation that happens to turn out and selecting a response as before, assuming that eventually the response made will be a good one. That is, it may matter *more* or *less* whether or not the system always makes the correct response if incorrect ones can be put right with little consequence.

However, there is another issue which is separate from our concern with how often the system makes the right or wrong response; although it will affect how often the system *does* get it wrong or right. This is the extent to which the outcomes to situations are inherently predictable when based upon the observable characteristics of those situations. If we are unable to correctly anticipate the outcome to a situation on this basis, then we require a means of supplementing the implicit prediction given us by situation-action rules. What we need to supplement it with is the ability to reason about the accuracy of the predictions on which we are to base action: to detect when naive prediction is inaccurate; to identify why the situation is unfolding in some other way; and to use that information to enhance the accuracy of subsequent predictions about the continuing development of the situation.

At this point, I would like to return to the issue of the need for responses to be fast. Discussion about action usually focusses on behaviour which directly achieves goals by bringing about changes in the world; however, actions for acquiring information also need to be carried out. In situations where there is continual change of an uncertain nature, information must be acquired before the need to act arises, so there is little time available. In addition, the opportunity for acquiring some kinds of information will be constrained to a limited time window. Not only, then, is the information to be apprehended continually changing and of an uncertain nature, it is not possible under all circumstances to apprehend *all* the information available.

Whether by circumstance or design - that is, be it haphazard or informed - the apprehension of information must be selective. As a result action and prediction must be based on only a limited subset of the available information. This diminishes the inherent predictive power of the observations made; because had we made some other (possible) observations we might have made other overriding predictions. I have worded this statement this way because I believe it strengthens the case against a simple reactive rule-based approach as being appropriate for all classes of uncertain, dynamic domains. Namely, had other observations been made, a different rule might have been deemed correct.

I believe it also strengthens the case for the appropriateness of a capability for reasoning about the accuracy of predictions upon which actions are to be based. Not least of all is the reason that predictions arising from more recent observations will be placed in the context of already existing expectations. Whether or not they entirely agree, posing a question over their accuracy, a more complete picture of the emerging situation will be built up than on the basis of those most recent observations alone.

Below, I describe a system, AUTODRIVE, which incorporates an initial attempt to implement the capabilities of predictive world modelling based on selective perception.

Dynamic World Modelling in AUTODRIVE

The AUTODRIVE system simulates the dynamic, uncertain domain of driving. A symbolic description of a situation and the changes that occur within it are presented as a series of 'snapshots'. Information about this world may be apprehended selectively by focussing on single domain entities over a sequence of such snapshots. The information is used to build up a dynamic world model for drivers of vehicles in the microworld. Every vehicle in the domain builds up its own unique dynamic world model based on the observations made from its own viewpoint. Of course, these viewpoints include drivers seeing each others' vehicles. The dynamic world model is a predictive model of how a situation will unfold over an eight second period. It is used to inform the actions of the vehicle driver using an integrated planning and dynamic decision-making approach which is not discussed further here but is described in detail in Wood (1993).

The Observation Mechanism

As each successive snapshot of the world is presented to a

vehicle driver, data about the current situation are made accessible to an Observation Mechanism (OM). This mechanism selects data pertaining to a single object[3] and passes it into a Short Term Information Store (STIS). During this selection process, the OM has access to information previously placed in the STIS. Objects which have previously been observed within a preceding interval of four seconds are ignored. The remaining observable objects are prioritised on an arbitrary preset rating scale according to class of object. This rating may be modified by requests received from the planning and decision-making processes, and processes for plan recognition, about the need to observe certain objects in particular. Once selected, an object is 'fixated' for 0.5 seconds (five time cycles in the current implementation).

The Short Term Information Store

Data describing each observed object, in a temporally qualified form, are stored in the STIS. The example below of data contained in the STIS describes a vehicle ahead of the driver, travelling at just under 30mph. Over a sequence of intervals the green car observes the red car travelling from a position 40.0 metres along the Lewes Road to 45.0 metres along the northern carriageway of the Lewes Road. The fixation start and end times are italicized.

stis,[greencar,[[vehicle,ahead],car1,[redcar,[lewes_rd6,lewes_rd, brighton],**40.0**,n,0.5,3.0,2.0]],*0.0*,0.1]
stis,[greencar,[[vehicle,ahead],car1,[redcar,[lewes_rd6,lewes_rd, brighton],**41.25**,n,0.5,3.0,2.0]],0.1,0.2]
stis,[greencar,[[vehicle,ahead],car1,[redcar,[lewes_rd6,lewes_rd, brighton],**42.50**,n,0.5,3.0,2.0]],0.2,0.3]
stis,[greencar,[[vehicle,ahead],car1,[redcar,[lewes_rd6,lewes_rd, brighton],**43.75**,n,0.5,3.0,2.0]],0.3,0.4]
stis,[greencar,[[vehicle,ahead],car1,[redcar,[lewes_rd6,lewes_rd, brighton],**45.0**,n,0.5,3.0,2.0]],0.4,*0.5*]

The Plan Lookahead Temporal World Model

The construction of the Plan Lookahead Temporal World Model (PLTWM) constitutes AUTODRIVE's dynamic world model. It is based on object location and vehicle descriptor data stored in the STIS. Once observed, an object's relative distance (fixed objects) or movements (vehicles) are modelled over a period of eight seconds. This model is represented by a sequence of data structures describing the object for each simulation cycle which will occur during the eight second period (that is, in this implementation eighty data structures are generated for each object) so that AUTODRIVE has access to the relevant projection descriptor for each interval during which decision-making will take place, over this period. An example projection for a fixed object would look like this:

projection,[red_car,[[sign, no_entry], sign5, [**54.0, 61.0**]], *9.6,9.7*]

This projects the relative distance, *eg* 54.0 metres and 61.0 metres, of a pair of no-entry signs from the viewpoint of the red car over a 0.1 second interval. The interval over which the object is predicted to be at the relative distance specified is given by the italicized figures, *eg* 9.6 - 9.7. The PLTWM bases its projection of the relative distance of fixed objects upon the known movements of the observing driver, decreasing by precisely the distance travelled by the driver per interval. If the driver intends to change the pattern of his behaviour during the projection period, the modeller takes this into account in generating the projected sequence.

Because other agents move independently of the driver viewing them, projecting the future locations of other vehicles involves ascertaining their speed and trajectory in order to calculate the absolute distance they will travel over the projection period. It also involves predicting likely behaviour changes. This firstly requires interpretation of the observed behaviour sequence. This is segmented, if required, at points where the vehicle is observed to have either changed direction, such as when reversing, or when, say, starting off after having been stationary for a period of time. A projection is based only on that behaviour observed since one of those segmentation points. The PLTWM bases its calculation of vehicle movement upon the degree of movement which takes place over the fixation interval (0.5 seconds) segmented in this way. However, before the future movements of the vehicle can be projected, expectations about changes in behaviour in satisfying situational constraints and the personal intentions of the driver must be identified. This process is described in the next section.

The Plan Recogniser

The Plan Recogniser (PR) uses the implicit description of action contained in the STIS to invoke hypotheses about the goal(s) a vehicle might be pursuing. The PR does not reason about the actions of vehicles *per se* but about their deviation from expectations. The initial projections of a vehicle constitute expectations; it is upon subsequent observation that deviations from expectations may be detected and hypotheses generated regarding the personal intentions of the driver.

The current speed and trajectory of the observed vehicle plus the expectation that it will attempt to travel at the speed limit provides a base line for generating expectations. The current situation[4] and the inferred personal intentions of the driver point up constraints on

[3]These information processing constraints do not apply to some categories of information, such as feedback to the driver on his own position, the curvature, gradient and width of the road and lane markings; that is, information that enables the driver to guide his vehicle. This decision was based on evidence that humans are able to acquire such data parafoveally, and is open to change.

this behaviour. For example, a vehicle travelling below the speed limit will be expected to increase speed at a reasonable rate of acceleration. From this, one may also calculate the distance the vehicle will have travelled in achieving this desired speed[5] assuming a reasonable rate of acceleration.

speed_constraint,[green_car, red_car, 1.3, 8.0]

Here the green_car hypothesises that the red car will behave in accordance with the constraint to drive at a speed of 1.3m over a 0.1 second cycle (approx 30mph) and is expected to have complied with that constraint by the time it has travelled the distance of 8 metres. At this point the vehicle will be anticipated to level off its speed and so the anticipated change in the vehicle's behaviour is also noted:

behaviour_change,[green_car, red_car, 8.0, 1.3, 0.0]

This describes the inferred intention of the red car to change its behaviour after it has travelled the distance of 8 metres in order to satisfy the constraint of driving at 1.3 metres per time cycle thenceforth. Because it is currently the red car's inferred intention to drive at the speed specified by the speed limit, this intention is also noted as the current immediate plan of the driver:

plan_constraint,[green_car, red_car, 1.3, 8.0]

All constraints are compared to ensure that always the strictest constraint applies (*eg* that which requires the greater rate of deceleration) through the constraint recording mechanism. For instance, a red traffic signal would lead the PR to anticipate the car ahead slowing down, and the speed constraint applying to that vehicle would be substituted with one reflecting this new information. Constraints might apply immediately, if appropriate, or at some point in the future as an anticipated behaviour change. In the latter case, the intention to change speed is noted as the vehicle's active plan_constraint.

For vehicles observed previously, earlier predictions are used to corroborate current expectations. A vehicle whose overall behaviour is not predicted on the basis of pre-existing or existing situational constraints, invokes reasoning based on the possible personal goals a driver may have and typical plans for achieving those goals. The PR attempts to link inferred behaviour (across the two observations) to a known plan. Supplementary evidence is sought to support this reasoning. For example, if a vehicle has pulled out and is decelerating, the goal of turning right (left in the USA!) may be hypothesised; therefore, a road junction on the right, more or less in the vicinity of where the driver will eventually come to a halt, would be good evidence in support of this hypothesis. Similarly, one might hypothesise that a vehicle rapidly slowing down may have noticed an obstacle or a red light ahead, previously unseen by our own driver; he would then seek supporting evidence of this too. The process of recognising other drivers' plans, then, makes further demands on the selective attention of the driver for confirmatory evidence of one kind or another. If the driver fails to find corroborating evidence of any kind, it is assumed that the vehicle will pursue its current course of action to its conclusion, modified to take account of current situational constraints.

The projection of the vehicle in the PLTWM is consequently modelled according to the constraints and behaviour changes identified. For example, the vehicle would be projected as complying with currently active constraints until the next predicted behaviour change is deemed to come into effect. Thenceforth the vehicle will be modelled as complying with the newly active set of constraints. This process continues for each predicted behaviour change until modelling is completed for the eight second period. The form of the PLTWM, then, is a sequence of projected snapshots describing every observed object and agent. It is sufficiently precise to identify interactions between those objects and agents and the constraints they impose on the viewer's own behaviour.

Selective Perception in AUTODRIVE

In the last section mention was made of directing attention to information for corroboration of plan recognition. Similar direction of attention takes place during AUTODRIVE's decision-making activities. In both cases an attention 'request' is generated and held alongside similar requests which, at the end of each dynamic world modelling and decision-making cycle, are passed to the Observation Mechanism (OM) described above. For example, here is an attention request by the driver of the green car to observe a direction sign. As it is as yet unobserved the *Data* it will provide is unknown. The request has a priority rating of 6 on an arbitrary ordinal scale of 1 to 10.

[attention_request,[green_car,[[sign,direction],*Data*,6.0]]]

Objects for which observation has been specifically requested in this way are awarded a higher than within class priority for observation so that, for example, a sought after junction would be prioritised above any other junction which just might come into view. However, another category, such as vehicles, would always receive a higher priority for attention than a junction.

[4]The current situation here being the *recorded* position of all objects and vehicles previously observed and modelled in the PLTWM corresponding to the *current time cycle*.

[5]Calculations are based on the standard equations of motion involving uniform acceleration: v = u+at; v*v = u*u+2as; s = t(u+v)/2.

Discussion

The AUTODRIVE system constitutes an initial attempt to incoporate dynamic world modelling, plan recognition, and a naive attentional mechanism into the planning process. As a vehicle for exploring these features AUTODRIVE has pointed up several limitations which are the focus of current and ongoing work.

Once generated, descriptions contained within the dynamic world model provide data about the unfolding situation, in the form of a sequence of projected snapshots. AUTODRIVE uses this information to identify important interactions between a vehicle driver and other agents and objects. Identifying such interactions leads to decisions about appropriate actions for the vehicle driver, of course. It also leads to modification or updating of the series of projections described in the dynamic world model, to reflect expectations about how other agents are likely to change their behaviour as they interact with the situation as it unfolds. A crucial feature of the dynamic world model, then, is that it replaces naive projections about agent behaviour with informed anticipation about how that behaviour may change as a consequence of that agent's hypothesised goals and complex interactions with the world. The dynamic world model then provides a more informed basis for anticipating correctly the likely behaviour of yet further agents interacting in this complex scenario, by identifying accurately the situation that will unfold.

The dynamic world model lacks the power of a full ATMS. However, just as a truly reactive system escapes the need for any form of situational representation; so it seems the form of the dynamic world model serves the needs and purposes of AUTODRIVE well. Dynamic world modelling provides a platform for responding to dynamic, uncertain domains by rendering those domains *certain enough* when this cannot be done purely on the basis of information apprehended at the moment of execution alone. In most cases this is because for such domains reacting comes too late - the need to act arises *before* the concrete evidence of perceptual information becomes available - and the greater the time lag between action and consequence, the greater the uncertainty about how the action will turn out.

Information available in the dynamic world model replaces the need for constant sensing of all available information in order to direct behaviour. Furthermore, the information requirements for plan recognition (as well as for planning and decision-making) can be used to direct perception in a selective manner, helping to overcome real-time constraints on sensing. The form in which an attentional mechanism is currently implemented in AUTODRIVE is crude. However, it succeeds in directing selective perception and serves as a basis for exploring these issues further. Work on improving the basis for attentional behaviour is ongoing, and is guided by the work of Hayes-Roth (1991) and Ogasawara, (1993), for example.

References

Ferguson, IA (1992) TouringMachines: An Architecture for Dynamic, Rational, Mobile Robots. Unpublished Thesis, Technical Report No 273, Computer Laboratory, University of Cambridge, Cambridge, CB2 3QG, UK.

Georgeff, MP (1987) 'Planning', Annual Review of Computer Science, **2**, *pp359-400.*

Georgeff, MP and Lansky, AL, (1987) 'Reactive Reasoning and Planning', *AAAI-6, Vol II, Seattle, WA, pp677-682.*

Hayes-Roth, B (1991) 'Making Intelligent Systems Adaptive'. In K VanLehn (Ed) *Architectures for Intelligence*, Hillsdale, NJ: Lawrence Erlbaum.

Sanborn, J & Hendler, J, (1988) 'A Model of Reaction for Planning in Dynamic Environments', *International Journal of Artificial Intelligence in Engineering,* **3(2)***, pp95-102.*

Ogasawara, GH, (1993) RALPH-MEA: A Real-Time, Decision-Theoretic Agent Architecture. Unpublished Thesis, Report No UCB/CSD 93/777, Computer Science Division, University of California, Berkeley, CA 94720.

Sanborn, J & Hendler, J, (1990) 'Knowledge Strata: Reactive planning with a multi-level architecture', Technical Report CS-TR-2564, University of Maryland, MD.

Wood, S (1993) *Planning and Decision-Making in Dynamic Domains*, Chichester: Ellis Horwood.