Investigating Agent Influence and Nested Other-Agent Behaviour

Robert Logie
Department of Informatics
Osaka Gakuin University
Osaka, Japan
rob@ogu-m.jp

Jon G. Hall and Kevin G. Waugh
Faculty of Mathematics, Computing and Technology
The Open University
Milton Keynes, England
J.G.Hall@open.ac.uk K.G.Waugh@open.ac.uk

***Abstract*—Stit theory, the notion of an agent's capacity for guaranteeing – or *s*-eeing *t*-o *it* – that a state of affairs arrives, provides a rich semantic framework for manipulating agent ability. However, the syntax of stit readily allows for nesting and such nested statements are not easily imbued with meaning. In this article, and in related work, we introduce and explore a partial logical characterisation of an alternative reading for stit based on a novel interpretation of agent *influence* that allows meaningful nested agent expressions.**

***Keywords*-Cooperative systems; Stochastic logic; Stochastic automata; Adaptive systems;**

I. Introduction, stit semantics

This article builds on the work of the authors [1] [2] [3] and from the first author's doctoral research [4] in providing theoretical underpinnings for leveraging agent interaction. Although we believe the work to be more general – there appears no characteristic dependence on reactive agents – we present the work using reactive agent for simplicity's sake.

The work extends *Stit theory* ([7]), an extension of situated first order logic in which an agent's effect on its world is considered.

Our extensions stem from the observation that, in the environment of a multi-agent system, it is not unreasonable to expect to be able to reason about one agent's effect on another; previously, however, there were difficulties in reconciling the notional independence of agents with that that they can influence each other. Indeed, simply by existing together, that agents influence each other is clear – a real-world agent that occupies a spatial location prevents another agent from occupying the same location; in stit theory, this is easily interpreted as influence. In this work, in addition, we allow one agent to have influence over the *internal state* of another: as might be the case when a contract is drawn up with clauses sanctioning any agent that does not fulfil its contractual obligations.

The extension is, we think, natural and provides a means for seeing to it that agents behave in a certain manner, and does not prevent the discovery of additional behavioural complexity for an agent group. Indeed, in the scheme we describe, *coaching agents* examine group behaviours with the intent of discovering *influence extending* joint actions. An example of extended influence through joint action is the weight that two agents can carry; another is the span of a "bridge" two can form together by joining.

We explore this notion of agent influence by developing a theory based on observed agent behaviour which we then implement in a system of coached reactive agents. We then introduce two influence operators, *leads to* and *may lead to* which we then use to outline elements of a logical characterisation of influence using the semantics of stit as a template.

We extend the use of stit semantics into the analysis of systems of simple reactive agents and will allow the analysis of agent influence in complex systems (which may have systems for obligation or sanction) based on the observation of their behaviour. Although the characterisation of our theory is incomplete, we present our investigations on some logical aspects of our influence based reading for stit.

The next section in this paper provides a short outline of stit theory and branching time, and describes the difficulties that standard stit has in addressing nested other agent behaviour. We briefly describe other areas where the notion of agent influence has been used and indicate where this differs from our work. Section III introduces our theory of influence and develops it to the point where we may describe nested agent behaviour. Section IV takes the theory into a practical setting by way of a simple system of coached reactive agents. The coaching agents use observations of agent behaviour in conjunction with the theory of influence to detect nested behaviour and to synthesise new behaviours which maximise agent influence. Section V partially explores logical characteristics of our influence theory using the stit framework as a template. Finally, Section VI summarises the paper and indicates future research areas based on agent influence.

II. Background

The first semantics for dealing with agency in terms of identifying actions with what they cause were laid out by Chellas [8]. This was taken in a number of directions with Belnap and Perloff [7], and, Horty and Belnap [9] focusing on the concept of *seeing to it*. This is usually cast against a background of possible worlds, a notion developed by Kripke, set in the branching time framework proposed by

Prior [10] and developed further by Thomason [11] [12]. Branching time represents the unsettled nature of the future by offering a number of possible paths forward and it represents the settled nature of the past by offering only a single path backwards. Each of these paths is called a *history* and points where a history divides are called *moments*. We use the notion of a branching temporal frame to describe this: a frame, $\mathcal{F}$, is a strict partial order $\langle Tree, < \rangle$ over $Tree$, a nonempty set of moments. A valuation function, v, maps propositions onto moment/history pairs thus $\mathcal{M} = \langle \mathcal{F}, v \rangle$.

At any moment, m, the future can divide into two or more histories: suppose h_1 and h_2 are distinct histories and that A holds only on h_1. Clearly m alone is insufficient for evaluating the future value of A, this most be done against a moment and history pair thus we say that A holds on m/h_1 but not on m/h_2. A proposition A is true when the valuation function indicates: $\mathcal{M}, m/h \models A$ iff $m/h \in v(A)$.

As agents makes choices, histories branch. As such, agent choices can be thought of as partitioning the set of possible futures passing through the moment where agents make that choice. Following Horty and Belnap [9] we assume a choice function which maps each agent, α on to a partition, $Choice^m_\alpha$, of all of the histories, $H_{(m)}$ through a moment m. Figure 1 shows a single moment, m, where an agent, α can choose between three actions J, K and L. Each action is distinguished by the histories it generates: L leads to a single history h_5, whereas both J and K lead to two each (presumably because of some phenomena hidden to us). The collection of histories at moment m for α, $Choice^m_\alpha$, is $\{\{h_1, h_2\}, \{h_3, h_4\}, \{h_5\}\}$.

More formally, the agent's choices divide the five possible histories into three equivalence classes according to its choice at m of J, K, or L. If α chooses K then it admits indeterminism as it does not control which of $\{h_3, h_4\}$ is followed, which are distinguishable because A holds only on h_4. We can therefore say that α choosing K cannot guarantee A. If α chooses L then it can guarantee that the future will evolve along h_5. In this case, α's choice of L ensures $\neg A$ will hold; α choosing L exerts influence over $\neg A$.

We blur the distinction between the choice open to an agent (the J, K and L) and the equivalence class of histories that a choice admits (the $\{h_1, h_2\}$, $\{h_3, h_4\}$ or $\{h_5\}$). A statement of the form $[\alpha\ stit\colon A]$ reads that an agent, α has the ability to see to it that A holds. In Figure 1 it can be seen that A holds on all of the histories emerging from J so that despite the indeterminism of $\{h_1, h_2\}$ the agent is able to guarantee A. This simple evaluation is a Chellas stit or $cstit$ that Horty and Belnap [9] state more formally as $Choice^m_\alpha(h)$:

$$\mathcal{M}, m/h \models [\alpha\ cstit\colon A] \text{ iff } \mathcal{M}, m/h' \models A \text{ for all } h' \in Choice^m_\alpha(h) \quad (1)$$

Given that $\models$ indicates the relation between an m/h index belonging to some model and the formulas true at that index, equation 1 says that A must hold on all histories emerging from some $Choice^m_\alpha$, α*'s* choice at moment m. Here we describe the notion of *settledness* in the branching time model. Horty and Belnap [9] indicate that settledness is historical necessity and liken it to the standard modal necessity operator $\Box$ and give an evaluation rule:

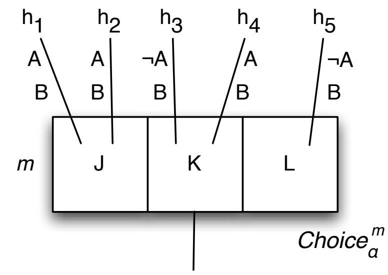


Figure 1. Example of both $[\alpha\ cstit\colon B]$ *and* $[\alpha\ dstit\colon A]$

$$\mathcal{M}, m/h \models \Box A \text{ iff } \mathcal{M}, m/h' \models A \text{ for all } h' \in H_{(m)} \quad (2)$$

This notion of necessity is incorporated into notion of *settledness* allowing us to describe propositions as being settled true or settled false at a moment and history pair m/h, following Horty and Belnap:

Definition 1: We define A as settled true at a moment, m, in a model $\mathcal{M}$ when $\mathcal{M}, m/h \models A$ for each h in $H_{(m)}$. Conversely, A as settled false at m when $\mathcal{M}, m/h \not\models A$ for each h in $H_{(m)}$.

The Chellas stit allows for but does not require that an agent be able to do otherwise and an alternative evaluation rule, the *deliberative* stit or dstit, adds this requirement, Horty and Belnap [9] state its evaluation as:

$$\begin{aligned}&\mathcal{M}, m/h \models [\alpha\ dstit\colon A] \text{ iff}\\&(1)\ \mathcal{M}, m/h' \models A \text{ for all } h' \in Choice^m_\alpha(h) \text{ and}\\&(2) \text{ there is some } h'' \in H_{(m)} \text{ where } \mathcal{M}, m/h' \not\models A \quad (3)\end{aligned}$$

Equation 3 says that in addition to Equation 1's requirement there must be at least one history, h'' in the set of all histories accessible to α at m, $H_{(m)}$, where A does not hold. The dstit reading forces the agent to make a deliberate choice Figure 1 shows that $\neg A$ holds on all of the histories emerging from L giving α a choice that guarantees $\neg A$ and providing the *negative condition* required for a deliberative stit and $[\alpha\ dstit\colon A]$ holds true at m/h_1. In contrast B holds on all histories $\{h_1, \ldots, h_5\}$ satisfying a Chellas stit but not a deliberative stit. The notion of an agent's choice at a *moment* may be extended by grouping moments horizontally across a tree, that is moments occurring at the same time, into a set known as an *instant*. This allows further refinement in the evaluation of stit expressions by applying temporal bounds. This broadens the scope of stit analysis to allow for chains of actions where one choice at a *witnessing* moment makes it possible for a later choice at the *achievement* moment in an achievement stit construct. The notion of witnessing moment is important to what follows so we state Horty and Belnap's [9] evaluation rule for an achievement stit or astit:

$$\begin{aligned}&\mathcal{M}, m/h \models [\alpha\ astit\colon A] \text{ iff there is a moment } w \text{ such that}\\&(1) \text{ for all } m'\ Choice^m_\alpha(h) \text{ equivalent to } m \text{ we have}\\&\quad \mathcal{M}, m'/h' \models A \text{ for all } h' \in H_{(m')} \text{ and}\\&(2) \text{ there is some moment } m'' \in i_{(m)} \text{ such that}\\&\quad w < m'' \text{ and } \mathcal{M}, m''/h'' \not\models A \text{ for some } h'' \in H_{(m'')}\end{aligned} \tag{4}$$

Equation 4 says that if $[\alpha\ stit\colon A]$ is true at m/h as a result of a choice by α at some prior *witnessing* moment then there are two requirements. The positive requirement, equation 4(1), is that as a result of $\alpha's$ earlier choice things have evolved in such a way that A is guaranteed now. The negative requirement, equation 4(2), is that at the witnessing moment A was not yet settled true so that $\alpha's$ choice at w had some real effect in bringing about A.

Definition 2: We define a witnessing moment for $[\alpha\ stit\colon A]$ as some moment w preceding m where a choice is made that allows α to see to it that A holds by a choice at m. Additionally, at moment w there must be a choice available which leads to a future where A is settled false.

A. Difficulties with stit

Consider a nested other agent stit expression of the form $[\beta\ stit\colon [\alpha\ stit\colon A]]$, i.e., β sees to it that α sees to it that A holds. This makes intuitive sense and Chellas [13] (also in Belnap et al. [14, page 275]) notes that it would be "...bizarre to deny that an agent should be able to see to it that another agent sees to something." Nested other agent expressions are syntactically well formed but their semantics fail when considered logically. When we say that $[\alpha\ stit\colon [\beta\ stit\colon A]]$ are we saying that α sees to it that β sees to it that A holds or are we saying that $\alpha's$ action makes it the case that β is *able* to see to it that A holds? If the former reading were true then α must exercise some influence over β. Belnap et al. [14, page 274], object to this reading on the grounds that it is inconsistent and justify this stance by demonstrating a contradiction. Rather than reproduce the proof here we represent a joint choice by two agents, α and β at a moment m_0 in Figure 2. I represents β seeing to it that A holds. If β acts in such a was as to see to it that A holds then this must be the case for *all* of $\alpha's$ choices because if this were not the case then α and β would not be independent agents. Figure 2 is described by Belnap et al. [14, pages 274–275] as representing a witnessing moment, m_0, the columns represent possible choices for α and the rows represent possible choices for β. If $[\alpha\ stit\colon [\beta\ stit\colon A]]$, then, where $I = [\beta\ stit\colon A]$, I must fill a choice column for α in m_0. But because I represents a stit by β, whenever I appears anywhere in a choice row for β in m_0 it must fill that row. So I must fill the entire diagram of m_0. If so then I is by definition settled true at m_0 and this contradicts the negative requirement, described in equation 4(2), that stit statements are never settled true at their witnessing moment.

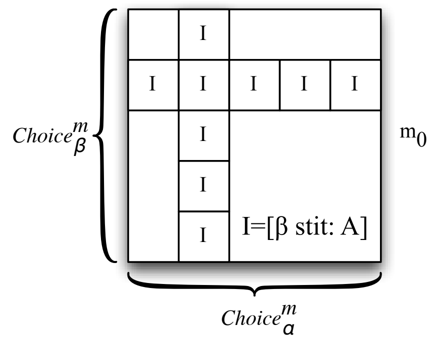


Figure 2. The logical impossibility of $[\alpha\ stit\colon [\beta\ stit\colon A]]$ (Adapted from [14, page 274])

Intuitively such statements make sense socially, α may have an obligation to β and in that way β may be thought of as seeing to it that α brings about A. Similarly β may be in a position to impose a sanction on α if α fails to bring about A. These allow for a meaningful interpretation of stit but they also bring burdens, simple reactive agents may have no notion of obligation and a system of sanctions adds an abstract choice – that of being sanctioned – for failing to bring about A. Belnap et al. [14, page 271] suggest a number of readings, including those above, that allow for a meaningful interpretation of nested other agent stit expressions but, as noted above, these may place additional requirements on agents making these approaches unsuitable for simple systems of reactive agents. When considered in the context of a society of agents, nested other agent constructs seem to imply that one agent may somehow control the other agent's choices thereby compromising its independence and agency.

The notion of influence introduced here is that one agent may influence another agent which, in turn, may use its influence to bring about A.

There have been other approaches to dealing with agent "influence". Ferber and Müller's multi agent based simulations (MABS) [15] formalism differs from this approach in that it aggregates agent actions to generate an effect on the environment. This approach, Michel [16] notes, does not readily model simultaneous actions and Weyns [17] indicates that it is limited to synchronous systems. Simonin et al. [18] extend Ferber and Müller's work into a formal design based on the B-Method. This work differs from these by treating both agents and the environment as "black boxes" and by operating on observation. Observations are used to generate hypotheses about agent abilities and these hypotheses are used with a theory of influence to generate behaviour patterns for reactive agents. Influence, in this setting, is treated as a logical property in an agent-centric stit like framework rather than something that is aggregated globally. An influence based reading of other agent expressions allows a meaningful interpretation without additional burdens, for example the cognitive ability to reason about sanctions and the extra system requirements for administering sanctions. Our influence reading allows us to extend the semantic reach of nested other agent expressions into the domain of simple reactive agents and possibly into the domain of emergent behaviour and the social systems that it is capable of building.

We consider cases of parallel action, where collocated agents act simultaneously, and serial action where agents are not necessarily collocated and do not act simultaneously. The *parallel* case is less complex because it allows agent choices to be grouped and possible futures mapped in the manner described by Horty [19, page 29]. The serial case is more complex and forms the larger part of our investigation. Where agent actions are separated temporally their choices cannot simply be grouped because the outcome of a sequence remains contingent until the final agent choice.

III. A THEORY OF INFLUENCE

We noted, in Section I, that our main interest is in simple reactive agents and the difficulty of allowing systems using these agents to improve their performance over time. How does such a simple agent evaluate its performance and, perhaps a greater problem, how may it alter its behaviour? We address these problems in two ways, we introduce an observer to the system and we allow this observer to synthesise new behaviours for agents. How far into the cognitive domain do we allow the observer to go? A cognitive observer presents subtle problems, consider a system as a state space with agents that are able to drive change in this state space. This state space will contain a number of "good" states and these in order to be achievable these good states must be reachable by at least one "route" from other states. A route, in this context, is a sequence of agent actions causing the environment to change from one state to another. A cognitive and intelligent observer may try to shape agent behaviour inappropriately because its intelligent choices may bias it towards certain behaviour patterns. This is not necessarily bad but it does not really fit with our notion of state space and routes, our notion is perhaps better characterised as a state space exploration and intelligently guided behaviour may leave parts of the state space unexplored.

This is where we introduce our notion of agent influence and we use this at the observer level Observer agents have no ability to alter the agent environment or to directly manipulate actor agents. They can generate new behaviour patterns that are placed in the environment for actor agents to acquire and this acquisition of behaviours should, over time, lead to actor agents becoming adapted so as to maximise their influence on their environment.

When an observer "knows" that a behaviour indicates that an agent has influence on some aspect of the environment then it may generate a new behaviour pattern for that agent, one that attempts to maximise the agent's use of that influential action. It is generally accepted that there are four conditions for an agent, α *knowing* that A [20, page 7]. α knows A if and only if (i) A is true; (ii) α believes that A; (iii) α has adequate evidence for A; and (iv) the relation between (i) and (iii) is not just accidental. Of the three conditions involving agents only (ii) is subjective and because of the we may quantify this subjectivity.

A. *Influencing influence*

The theory of influence described here follows Milner's [21] observation that "the behaviour of a system is exactly what is observable". In addition, our theory admits that in noisy environments and in cases where agent ability is contingent and not fully understood then evidence may appear to be inconsistent. Where behaviour that involves two agents is observed then it may be said that $[\alpha\ \textit{influences}\colon [\beta\ \textit{influences}\colon A]]$ without compromising the agency of either party.

Although stit semantics present difficulties in other agent settings they provide the foundation for our theory of influence. The simple reason for this is that stit semantics are rich and expressive and provide a good template for influence. Stit expresses agent ability by characterising the partitioning of possible futures according to agent choices. If an agent has unambiguous ability to bring about A then observations of its behaviour would be similar to Figure 1 where at least one choice guarantees A. We introduce a notion of *strict* stit here and use this as a basis for differentiating our influence based reading from the standard *strict* reading, we observe that:

Observation 1: Standard stit expressions have strict requirements for the truth values of propositions, if a proposition does not hold following an agent choice then a related stit expression will be falsified. Influence has weaker requirements for the truth values of propositions, a proposition not holding following an agent choice does not necessarily mean that that choice has no influence over the proposition.

The strict approach carries the implication that one agent has control rather than influence over the other agent or agents involved in a complex behaviour. It is, thus, the strict reading that presents difficulties and we intend our influence reading to be as semantically close as possible to strict stit and also to maintain the agency of all agents involved in a complex behaviour. Suppose that α*'s* ability to bring about A is contingent on another agent's choice, what would be observed then?

Consider choice K in Figure 1, it can be seen that if α chooses K, written α/K, then this leads to two possible histories, $\{h_3, h_4\}$ with A holding on one but not the other. α, it seems, can not see to it that A holds by choosing K.

Inspecting Figure 1 shows that there are histories with A and histories with $\neg A$, this means that A is neither identically *true* nor identically *false* and may, potentially, be under the influence of an agent or agents. This allows for a hypothesis that α/K has influence over A and historical data may be inspected for supporting evidence. There are three forms of evidence that may be observed for a hypothesis that α/K has influence over A, these are:

Observation 2: We observe that Positive evidence for a hypothesis that α/K has influence over A is an instance where A is observed following α/K.

Observation 3: We observe that negative evidence for the same hypothesis is an instance where $\neg A$ follows a choice from the $\neg K$ partition.

Observation 4: We observe that counter evidence for the same hypothesis is an instance where $\neg A$ follows α/K.

First, assume that α/K has unambiguous influence over A in a noise free environment. Then, over a number of observations, there would be no instances of counter evidence, a number of instances of positive evidence and at least one instance of negative evidence. The number of instances of negative evidence beyond one is immaterial, it simply serves as a flag that A is not constant. Representing observations of

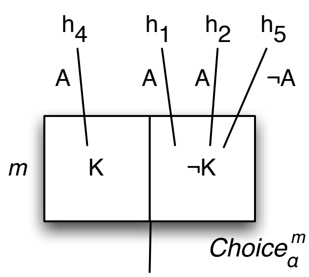


Figure 3. α/K has unambiguous influence

noiseless, unambiguous influence as a branching time choice gives Figure 3 where α/K clearly leads to h_4 where A holds. This satisfies the *strict* reading for both cstit and dstit that was outlined in observation 1. In a noisy environment there may be cases where α/K does not lead to A and such counter evidence would manifest itself as an additional history from the K partition at $Choice^m_\alpha$ where $\neg A$ holds. This means that $Choice^m_\alpha$ no longer satisfies strict stit evaluation rules because α has no choice by which it may guarantee A.

B. Extended influence and gateways

A single agent operating in isolation from other agents may be able to influence its environment in a number of ways. We refer to the set of world states that this agent can bring about by its influence as its *domain of influence* and for this single agent, the *single agent domain of influence* and in a sufficiently complex world this will be a subset of all possible world states.

Definition 3: Given an environment, e, that has a set, E, of possible states and a single agent, α, situated in e then we define α*'s* single agent domain of influence as the set of states , $s \in E$, accessible to α as a result its influence, $S^0_\alpha \subsetneq E$.

Here the 0 suffix indicates that we are considering single agent influence with no contribution from other agents. We extend definition 3 to the notion of extended influence by considering the intermediate stage of multiple, independent agents. If an environment contains a number of agents and these agents are independent of each other then we may aggregate their individual influence domains to give a set of reachable states for that world.

Definition 4: Given an environment, e, that has a set, E, of possible states and a set of agents, $G = \{\alpha, \beta, \ldots, \omega\}$, situated in e and with a set of possibly overlapping single agent domains of influence $\{S^0_\alpha, S^0_\beta, \ldots, S^0_\omega\}$. We define $S^0_G = \bigcup_{\alpha \in G} S^0_\alpha$ as the aggregate set of individually reachable states for that group of agents.

The 0 suffix, as in the single agent case, indicates that although we are considering a group of agents we are confining this consideration to single agent influence. Moving on to the two agent case we consider agents α and β situated in an environment where they can influence each other and assume that by working together – either in parallel or by serial cooperation – that at least one agent is able to reach world states that were unreachable to it individually. We must be careful here and emphasise that *is able to reach* means *is able to reach as a result of its action* so that we do not attribute α*'s* perceiving a state brought about by β as α*'s* reaching that state. This is what we term *extended influence* and we shall define this for two agents before considering a more general definition.

Definition 5: Given an environment, as above, a set of two situated agents, $G = \{\alpha, \beta\}$ and S^0_G, an aggregate set of individually reachable states. We say that S^1_α describes the set of states reachable by α operating in conjunction with one other agent, in this case β. α exhibits extended influence when $S^1_\alpha \backslash S^0_G \neq \emptyset$.

then by working together they may reach world states lying outside of their individual single agent domains. In such cases we state that an agent's individual influence has been *extended* into another domain and, for convenience, refer to these as two agent domains, three agent domains and so on.

Before considering the general case we note that influence domains may be notionally nested relative to a given agent's individual ability. This is illustrated in Figure 4 that may be thought of as representing two agents, α and β "meeting" in their environment, at this point α may give β some token or tool. This is a "gateway action" that allows β to influence a new world state which is contained in the two agent influence domain. It may be that α does not choose to the gateway action and in this case β will remain in the single agent influence domain. If two agents acting in together simultaneously can bring about an individually unreachable state then each agent's influence extends the other agent's ability and does so without either agent *seeing to it* that the other agent does something. Other agent actions, whether

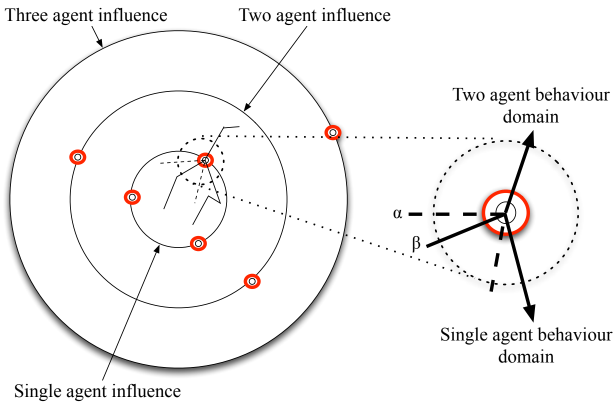


Figure 4. Nested influence domains relative to an agent

these occur simultaneously or in series, that allow an agent to move between domains are characterised as "gateway" actions notionally representing gateways between influence domains, we observe that:

Observation 5: Where one agent, α acts in such a manner as to allow some other agent, β do bring about a world state that was previously inaccessible to β then we observe that α*'s* action is a *gateway* action that allows β to extend its influence into a higher influence domain.

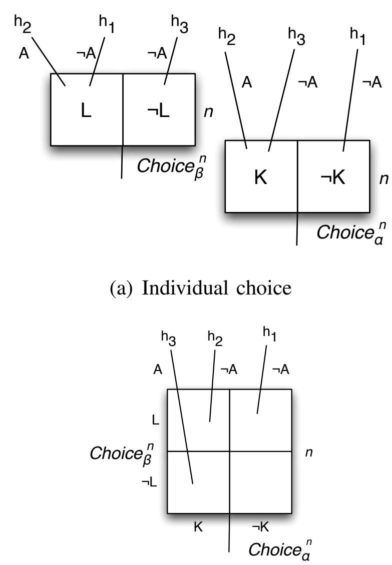
This representation highlights two interesting properties. If agents are to jointly bring about a proposition then the state space must contain that proposition. This may seem like an obvious requirement but it provides a clear representation of the deontic *ought implies can* [19] and one that may be applied to learning and adaptive systems. The second property is that there may be few *gateways*, if a state space is to be searched then locating and analysing gateways will require some form of guided searching.

Returning to our definition of extended influence and bearing in mind the notion of domain nesting we consider the more general case. Nesting brings obvious problems, if two agents acted alternately each extending the other's behaviour causing the world to cycle between individually unattainable states then we may see infinite nesting of repeated states.

Definition 6: Given an environment, as above, and a set of agents, $G = \{\alpha, \beta, \dots, \omega\}$, situated in e then for a single agent $\alpha \in G$ we define the n^{th} level of nested influence for α as $S^n_\alpha = \bigcup_{i \le n} S^i_\alpha$ and where α has extended influence at level n then $S^n_\alpha \setminus S^{n-1}_\alpha \neq \emptyset$.

Definition 6 allows us to complement levels so as to prevent repeated states from leading to infinite chains. This defines extended influence for a single agent in a world where other agent influence is implicit and may readily be extended to groups of agents as necessary.

Definitions 3 to 6 have viewed influence in a set theoretic manner and we now return to the earlier branching time representation of agent action. The extended influence part of ability may be characterised as a variation in the mapping of histories to the choice partition seen by an agent at a moment. Considering, briefly, the case of a parallel collocated action where two agents act simultaneously and in doing so extend the ability of one or both agents. This may be represented by equations 5 and 6 below. These indicate that a joint action has an effect on the distribution of histories causing it to differ from the distribution in an individual agents choice partitioning. Consider the two agents of Figure 5(a), α and



(a) Individual choice

(b) Parallel choice

Figure 5. Individual and parallel choice

β, with $Choice^m_\alpha = \{K, \neg K\} = \{\{h_2, h_3\}, \{h_1\}\}$ and $Choice^m_\beta = \{L, \neg L\} = \{\{h_1, h_2\}, \{h_3\}\}$. Note that neither α nor $beta$ influences A individually. Only when combined, so that the equivalence classes are further partitioned (see Figure 5(b)) by their joint action is influence evident, as then $Choice^m_{\alpha\|\beta} = \{(L, K), (L, \neg K), (\neg L, K), (\neg L, \neg K)\} = \{\{h_2\}, \{h_1\}, \{h_3\}, \emptyset\}$.

Horty and Belnap [9] use the notation $Choice^m_\alpha(h)$ to represent the particular possible choice by α at moment m that contains history h. This choice may contain other histories and we extend the notation by prefixing it with an h, $h.Choice^m_\alpha(h)$, that reads *some history belonging to the choice partition containing* h. We also combine agents so as to allow a joint choice, $Choice^m_{\alpha\|\beta}$ indicates that the two agents α and β are acting in parallel and exercise their independent choices at moment m.

This allows us to state that in general, a necessary condition for joint action being more influential at a moment m is that:

$$\exists h.Choice^m_{\alpha\|\beta}(h) \subsetneq Choice^m_\alpha(h) \tag{5}$$

or

$$\exists h.Choice^m_{\alpha\|\beta}(h) \subsetneq Choice^m_\beta(h) \tag{6}$$

Informally, joint action by α and β at moment m alters the history or histories passing through some choice partition so that for the joint action the valuations of propositions on histories, the number of histories or both will differ from those of the individual agents making the same choices independently. This allows for a measurement of influence, if the set of world states accessible to α and β acting jointly contains states that are not in the sets of accessible states available to the individual agents. "Measurement" is not linear in that we are considering the number of accessible states, it is based on set membership and we consider the joint reachability of states that are individually unreachable as *more influential*. In this example the equations identify h_6 where A holds. Agent influence in a collocated and cooperative action may be thought of as being commutative, α cooperating with β is the same as β cooperating with α. Sequential influence is not commutative and may require an ordering of actions. For example, a β type agent requires some token to allow it to bring about A. If a α type agent acts so as to give the β agent this token then it must do so *before* β chooses its action. This is illustrated by Figure 6 where moments, that are agent choices local to a particular history, potentially occurring at the same time are grouped into sets of *instants*. Here $I_1 = \{m_1, m_2\}$ and $I_0 = \{m_0\}$. If α/L at instant I_1 and this is followed, at instant I_1, by β/K then this guarantees A. Conversely, if $\alpha/\neg L$ I_0 then β has no choice available at $m_2 \in I_1$ that guarantees A. Where we write $Choice^m_{\alpha;\beta}$ we say that β executes an independent choice at m and that this choice has been preceded by a choice by α at some earlier moment. For the sequential case, a necessary condition is that

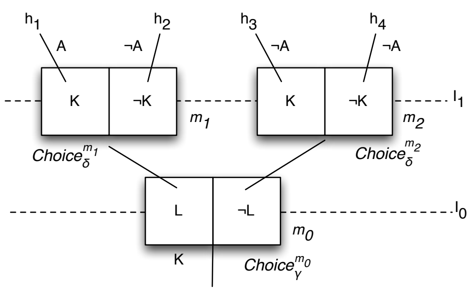


Figure 6. Sequential choice

$$\exists h.Choice^m_{\alpha;\beta}(h) \subsetneq Choice^m_\beta(h) \tag{7}$$

Informally this reads that the mapping of histories for a given choice by β at moment m when preceded by α*'s* action differs from what it would have been in the absence of α*'s* action. Intuitively, α*'s* action plays a part in refining the distribution of histories in β*'s* choice partitioning so as to either remove uncertainty or add new histories and potentially extend β*'s* ability.

IV. Observing and reasoning about influence

In observations 2, 3 and 4 we outlined how evidence presents itself and we note that evidence is countable, a coaching agent may tally how many pieces of evidence it has gathered for each evidence class. The coaching agent has only these data to work with so we use these tallies to generate indices for each hypothesis allowing a coach to rank the hypotheses in its database. The first measure adopted was the ratio between positive and counter evidence and we call this the *P:C* ratio for a hypothesis.

Definition 7: Given a hypothesis that an agent action leads to a proposition holding, a tally of instances of positive evidence for this, *P*, and a tally of counter evidence, *C*, we define the *P:C* ratio for this hypothesis as the ratio of these observed tallies.

Early, and very simple, experiments were carried out with a single class of agents each possessing identical actions:

Definition 8: We define an agent class as a set of agents possessing identical sets of actions.

These experiments indicated that the *P:C* ratio was a useful measure but when noise was introduced or where nested behaviours were being explored the *P:C* ratio alone was not sufficient and we sought another way of using the observed data to rank hypotheses.

Definition 9: Given a hypothesis that an agent action leads to a proposition holding, a tally of instances of positive evidence for this, *P*, and a tally of counter evidence, *C*, we define the *P-C* value for this hypothesis as the difference between the positive evidence tally and the counter evidence tally.

We adopted the *P-C* value as an additional metric allowing coaching agents to filter groups of hypotheses with similar *P:C* ratios. Intuitively a larger *P-C* value for a hypothesis indicates that there are more incidences of it showing influence than not and is, possibly, worth considering synthesising a behaviour based on this hypothesis and distributing it in the agent environment. The negative evidence value was not used here because, as indicated above, it simply acts as an indication that a proposition is changeable, if there was no negative evidence then there would be no hypothesis to consider.

If the environment were noisy and this noise had some effect on the value of A then observations would be unlike those illustrated in Figure 3. Instances of counter evidence may be observed but since observations already indicate that α/K has some influence over A we use the *P:C* ratio (in experiments this was *positive / (positive + counter)* so as to prevent divide by zero errors) to indicate the relative strength of each item of evidence. Clearly if the *P:C* ratio is 1 then influence is unambiguous, if the *P:C* ratio <1 then either the agent has influence and there is some interference or the agent has no influence and something else is changing A. The *P:C* ratio provides a value that allows the ranking of hypotheses and the tracking of the effect of any changes in behaviour. If, for example, α/K has influence over A is contingent on β/L then increasing the incidence of β/L will result in stronger evidence for α/K having influence over A.

If α*'s* influence over A is contingent on a choice by β, either at some earlier time or at the same time, then noisy evidence would be observed but the *P:C* ratio for α*'s* influence will also incorporate the influence of β*'s* action. This other agent influence, when combined with noise, will exhibit a lower *P:C* ratio than unambiguous influence. If α has no influence over A then the *P:C* ratio will be dominated by instances of counter evidence and will be low. This gives three *bands* of evidence, one for unambiguous influence, one for other agent influence and one for no influence. A *P:C* ratio band may contain a number of hypotheses for agent influence but which of these are interesting and worth investigating further? This was investigated by a series of experiments carried out in a simple agent system. In this system a number of reactive *actor* agents of the same type or class were observed by dedicated *coaching* agents. These coaching agents aggregate observations of agent behaviour – an agent's pre-action perception of its environment or *precepts* its action choice and its post-action perception of its environment or *postcepts* – using them both to identify agent influence and as a foundation for synthesising behaviours that will maximise agent influence individually and collectively.

A. Finding unambiguous influence in a noisy environment

Coaching agents were able to observe actor agents by examining history data that actors left in the environment. Based on these observed data coaching agents generated hypotheses for actor agent influence and seed the environment with new behaviour patterns based on what were considered good hypotheses. Identifying good hypotheses solely by observation presented some interesting problems. The notion was that if coaching agent hypotheses correctly identified agent influence and biased agents towards exercising that influence then this would be reflected in later observations.

Simple experiments in an extremely noisy environment demonstrated that ranked *P:C* ratios tended to fall into bands [4]. Unambiguous agent influence hypotheses were grouped at the top with hypotheses containing influence that was potentially contingent on other agents falling into a group separated from and below the top hypotheses. A third band containing "poor" hypotheses separated out at the bottom of the ranking. The banding of *P:C* ratios is, intuitively, heavily dependent on factors such as environmental noise and, in cases of other agent influence, the number of instances of β selecting the appropriate action to enable α to bring about A. The banding of hypotheses indicated that unambiguous influence was easily detected but ranking by *P:C* ratio alone was insufficient to allow the identification of potential other agent candidate hypotheses.

B. Finding other agent influence in a noisy environment

At this point evidence for hypotheses is considered solely on the basis of metrics, described above, generated from positive and counter evidence tallies. The *P:C* ratio alone proved to be insufficient and although considering it in conjunction with the *P-C* value yielded some improvement but a noisy, multi agent environment still presented difficulties. Further experiments led to a system of prefixing the hypotheses with agent precepts. This allows an agent to use what it perceives of its world as a filter to refine how it may use its ability. The intuition behind this is twofold, firstly the hypotheses would embody the notion that, neglecting effects of noise and other agents, an agent's ability is truly contingent on its current state and the states that are reachable from that state. The second intuition was that this would provide finer means of filtering hypotheses. Rather than a general hypothesis that α choosing action K leads to A holding, which we write as $\alpha/K \rightsquigarrow A$, with a single set of evidence tallies a hypothesis is grouped with a set of precepts, P, then α*'s* choosing K leads to A and we write this $\alpha^P/K \rightsquigarrow A$. Using precepts as a filter allows the introduction of preconditions for an agent's ability and in doing so we gain the means to analyse sequences of actions. Recalling the deontic notion briefly mentioned in Section III-B, this precept filtering gives a *situationist* dimension, as described by Hansson [22], which allows for agent ability to be categorised by the agent's situation. Coaching agents are thus able to select the most influential hypothesis for a given set of agent precepts. These explorations of simple systems, in conjunction with work on a partial logical characterisation of influence operators, provided sufficient information to build a coaching agent for use in a simple world.

C. Experimental observations

The single agent class experiments, outlined in Section IV, were extended to use two agent classes and provide a system that would test the practical application of the influence theory outlined above. These were based on the idea of incrementing an accumulator and brought the possibility of agents being able to extend their influence by repeated actions. Actor agents were, initially, simple stochastic agents with a fixed set of actions from which they would choose at random, each of the choices was weighted evenly so as to give a flat probability distribution. Coaching agents would be able to observe agent influence but would be unable to see that an influence was bringing the system closer to a specific value that was considered as being a goal state. When coaching agents observed influence they generated new behaviours that are probability distributions biased towards a particular influential agent choice. The use of a numerical accumulator is similar to the notion of bridge building, agents bridging a gap may exhibit extended influence by their being able to move further across a gap but the results of that extended influence will not be evident until the gap has been successfully bridged.

Initial validation experiments with a single agent class showed that both in noiseless and noisy settings agent coaching led to agents behaviour rapidly becoming biased towards incrementing the accumulator. The system was extended to two agent classes with one class, α being able to increment an accumulator from zero and the other class, β only able to increment non zero accumulators. This provided a simple case of nested other agent ability, $[\alpha\ \mathit{influences}\colon [\beta\ \mathit{influences}\colon A]]$. It was observed that both agent classes were coached so as to maximise their influence and the environment was extended so as to contain a number of agents of each class and a number of accumulators. Coaching, again, had agents rapidly maximising their influence and as agents moved to other accumulators they were quickly able to increment accumulators to target values. An interesting point in this experiment is that "good" aggregate behaviour patterns involve very little observable action from α class agents. The number of instances of α class agents incrementing an accumulator from its zero state will be small when compared with the number of instances of β class agent influence resulting from repeated incrementation. The frequent influence of β class agents ought to be easy to spot but what of the relatively small contribution of α class agents? This is a contribution that is so small that it may be readily swamped in a noisy environment. Evidence is based on observation and in a noisy or uncertain environment and, as noted in Section III-A, we may expect to observe some counter evidence to an agent's ability. Observed counter evidence does not necessarily mean that an agent has no influence. It may be that $\alpha's$ ability is contingent on another agent or it may be that another agent acted simultaneously so as to counter $\alpha's$ choice. Experimental data from an agent test are presented in Table I. The first column, ID, is simply an ID tag for a particular hypothesis. The precepts column contains data from an agent's precepts and the two digits indicate the presence of an accumulator and if the accumulator is zero respectively. The third column describes the hypothesis in question, for example, $\beta/5 \rightsquigarrow AC : 3$ says that the hypothesis is that a β class agent choosing action 5 leads to transition type 3 being observed in an accumulator. Type 3 transitions indicate that the accumulator incremented and type 1 transitions indicate that the accumulator became non zero. The remaining columns list the *P:C* and *P-C* values for observations of these hypotheses. Since the coaching agent had no cognitive or reasoning abilities, its operation was based solely on observation, generation of hypotheses and seeding of hypotheses selected by the *P:C* and *P-C* metrics. Table I is divided into three horizontal groups, the top group lists the hypotheses chosen as seeds for generated agent behaviours. These are the most influential behaviours for each agent class and each transition or change in the environment. Behaviour ID 1, $\alpha/5 \rightsquigarrow AC : 1$ has very poor *P:C* and *P-C* values but it remains the most influential behaviour for α agents leading to transition 1 and because of this it is considered suitable for seeding. The groups below show the top six hypotheses in the database ordered by *P:C* and *P-C* values.

Table I
HYPOTHESIS DATA, NOISY ENVIRONMENT

ID	Precepts	Hypothesis	P:C	P-C
Generated behaviours				
0	10	$\beta/5 \rightsquigarrow AC : 3$	0.884134	13122
8	10	$\alpha/4 \rightsquigarrow AC : 3$	0.627907	11
1	11	$\alpha/5 \rightsquigarrow AC : 1$	0.00050045	-9981
Hypothesis database, P:C ordering				
0	10	$\beta/5 \rightsquigarrow AC : 3$	0.884134	13122
8	10	$\alpha/4 \rightsquigarrow AC : 3$	0.627907	11
6	10	$\alpha/6 \rightsquigarrow AC : 3$	0.580645	5
9	10	$\alpha/2 \rightsquigarrow AC : 3$	0.555556	4
10	10	$\alpha/0 \rightsquigarrow AC : 3$	0.53125	2
3	10	$\alpha/3 \rightsquigarrow AC : 3$	0.0597125	-10780
Hypothesis database, P-C ordering				
0	10	$\beta/5 \rightsquigarrow AC : 3$	0.884134	13122
8	10	$\alpha/4 \rightsquigarrow AC : 3$	0.627907	11
6	10	$\alpha/6 \rightsquigarrow AC : 3$	0.580645	5
9	10	$\alpha/2 \rightsquigarrow AC : 3$	0.555556	4
10	10	$\alpha/0 \rightsquigarrow AC : 3$	0.53125	2
7	10	$\beta/1 \rightsquigarrow AC : 3$	0.00885609	-1331

Our intuitive knowledge of the problem indicates that α class agents have an influential action – the tipping of an accumulator from zero to some non zero value – and that the occurrence of this action would be very small. The *P:C* and *P-C* orderings worked as expected, the most influential actions rose to the top of the ordered tables. The *P-C* value is intended as a filter to offer further differentiation of hypotheses with very similar *P:C* ratios. However, the gateway action, α*'s* initialisation of the accumulator to a non zero value, does not appear in either ordering. Earlier work had led us to consider the nature of hypotheses which, until then, had been global, that is they took no account of the agent's immediate environment and were based solely on an agent choice or action. This meant that coaching agents had an implicit assumption that, for example, a hypothesis that $\beta/5 \rightsquigarrow AC : 3$ would be applicable in all circumstances. This is clearly far too coarse, the transition from zero to non zero is only going to be occur if the accumulator is zero before the agent makes its choice of action. This led us to generate "prefixed" hypotheses. The prefixes were, as described above, the agent's pre-choice precepts. Returning to the generated behaviours section of Table I and using precepts we see that given precepts 10 (the presence of a non zero accumulator) the hypothesis $\beta/5 \rightsquigarrow AC : 3$ exhibits the greatest influence for β class agents.

The most influential hypothesis for α class agents given the same precepts is $\alpha/4 \rightsquigarrow AC : 3$. This illustrates the overwhelming effects of noise generated by the actions of other agents. α class agents have no influence here but if an α class agent is collocated with a β class agent and the β agent does increment the accumulator then that change will appear in the α class agent's postcepts whatever action it selects. The hypothesis ranking shows that the β class behaviour is clearly more influential and this is, indeed, correct. For a different set of precepts, the presence of a zero accumulator, we have only observed a small number of influential actions but for this set of precepts it is the most influential action observed so is a candidate for seeding a new behaviour for α class agents.

Agents were built so as to accommodate multiple behaviours. These are represented as sets of weightings for each of the agent's possible choices or actions. The default behaviour has a flat weighting making each of its actions equally likely by random selection. Acquired behaviours are biased towards a particular action and the strength of this bias depends on the strength of the coaching agent's evidence for that action being influential. Each acquired behaviour has an associated percept pattern that is used to select the appropriate bias to apply to action selection.

We noted, above, that agents have a fixed fixed set of abilities, they are unable to acquire new choices or actions but are able to acquire new preferences or weightings for the selection of these abilities. Combining these fixed abilities with precept prefixed hypotheses allows us to consider actor agents as a mapping of precepts on to preferred behaviour patterns. This is illustrated by the schematic agent architecture of Figure 7 where an agent consists of a collection of behaviours selected by some precept filtering mechanism. Conceptually actor agents are a collection of finite state machines each with a weighted stochastic transition selection mechanism.

The agent, illustrated in Figure 7, has six choices available to it, $Choice_{agent} = \{K, L, M, N, O, P\}$. Each of these choices represent some action available to the agent, the choice of action is individual and independent of other agents but it may be in concert with other agents or it may be a null action. The set of choices remains constant throughout an agent's life, it cannot acquire new choices

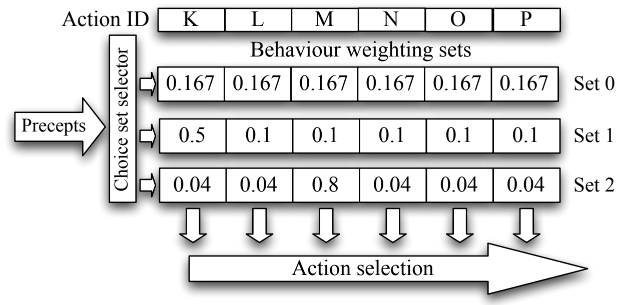


Figure 7. Agent internals – behaviour stack holding three selectable behaviour patterns for actions K to P.

and it cannot discard any of its current set of choices. The agent's choice set is overlaid by a set of weightings with each weighting assigning a preference for a single choice. Choice set 0, in Figure 7, gives each of the elements of $Choices$ an equal weighting, this is the agent's default state and it has no individual behaviour characteristics, each of its actions has an equal chance of being chosen. Behaviours 1 and 2 are acquired behaviours that have been synthesised by a coaching agent. Behaviour 1 is biased towards action K and behaviour 2 has a stronger bias towards action M. A behaviour pattern will not be allowed to become an absolute choice, we maintain a very small stochastic element so as to prevent the system becoming trapped by false hypotheses and to allow agents to adapt to changes in their environment. When the actor agent is operating it uses its precepts to select an appropriate behaviour pattern, it there is no match then it selects the default pattern It then generates a random number, checks this against the weightings in the selected choice set and executes the its choice.

Coaching agents generate a database of observations during their operation. This database contains hypothesis data, as listed in Table I, and is structured so as to reflect hypotheses overlaid on branching time. The notion of branching time brings difficulties, it is unbounded in nature and its relentless forwards branching is clearly not suited to bounded computational systems. To address this difficulty we admit the use of loops in the branching time structure. Where an agent's action leads to a particular state then we link that action to a "state bucket" that contains a chain of hypotheses for agent actions given that state as a precept. If a hypothesis leads to the same state then the link from that hypothesis goes to the state bucket at the root of its chain. If there are multiple versions of the same $agent/action$ hypothesis in a chain each will lead to a unique next state. This brings the database structure closer to the notion of *leads to* and *may lead to* relations that are embodied in our hypotheses and discussed in some detail in Section V. Figure 8 illustrates this, the boxes on the left hand side are "state

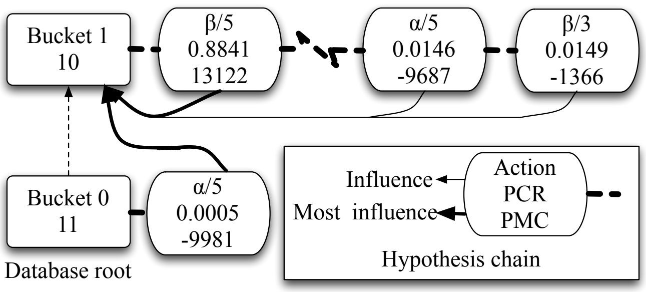


Figure 8. Partial coaching hypothesis database from experimental data

buckets" that represent the agent precept prefixes described in Section IV-B. Precepts are shown as two binary flags that indicate presence and state of an accumulator. Bucket 1 holds a list of influential actions when agents perceive a non zero accumulator and bucket 0 holds a list of influential actions where agents perceived a zero accumulator. The dotted lines leading right from the state buckets indicate hypothesis chains, from bucket 1 it can be seen that $\beta/5$ leads to a non zero accumulator (influence is identified by a change in accumulator value) with a *P:C* ratio of 0.8841 and a *P-C* value of 13122. Further along that chain we see that $\alpha/5$ leads to a non zero accumulator with a *P:C* ratio of 0.0146 and a *P-C* value of -9687. The most influential hypotheses in a chain are represented by a heavier line from the hypothesis entry to the finishing state bucket.

By inspecting the data it may be seen that from state 11 there is no influential action available to β class agents and that, despite the very low *P:C* ratio and poor *P-C* value, the $\alpha/5$ hypothesis is the most influential that we have observed from a world state where the accumulator is zero. In inspecting the data of Figure 8 we apply human reasoning and infer that if a system is in state 11 then the only way to activate β class agent influence is for α class agents to take the system out of state 11. A simple coaching agent only "sees" that there are some actions that, from observed evidence, are more likely to be influential than others and it seeds agents with these behaviours. The system may behave as intended but does so as a result of random interaction between independent behaviours, the coaching operation has simply increased the incidence of influential behaviours. In order to generate patterns of behaviour that cause agents to behave as required the coaching agent needs some means to reason about individual behaviours in a way that allows for sequencing, combination and substitution. This may also allow the coaching system to coordinate agent behaviours, perhaps by the use of null actions that Halbwach [23] indicates is standard practice in synchronous reactive systems programming. In order to do this we need to characterise agent influence logically using a framework as close to that of standard stit as possible. This will indicate what a coaching agent may safely infer from observed sequences of behaviour and guide how it may synthesise more complex aggregate behaviours. In the following section we outline some aspects of this work.

V. Logical aspects of influence

In order to explore logical properties of influence we investigate the *leads to* operator introduced in Section IV-B. Our theory is based on observations that may provide evidence of influence. The first, $\rightsquigarrow$ a *leads to* operator, may be considered as being similar to a stit operator. For $\rightsquigarrow$ to hold we must observe evidence of ability and consistent results. Although we think of the *leads to* as stit-like it differs from stit in that it is based entirely on observation and does not necessarily represent a complete characterisation of an agent's choices or exploration of its world. The similarity allows the characterisation of influence in a modal setting, the $\rightsquigarrow$ operator is very close to standard stit and in an ideal setting may be equivalent. If an agent, β*'s*, ability is contingent, perhaps on another agent's choice that must occur before or at the same moment as β*'s* choice then we may see inconsistent evidence for β*'s* ability. To allow for this we introduce a *may lead to* operator that we write as $\Diamond\!\!\rightsquigarrow$. This has weaker evidence requirements, a hypothesis based on the notion that an agent choice *may lead to* S allows for counter evidence. The presence of counter evidence is taken as an indication that the agent's ability is not fully understood and may be dependent on other agents or environment factors. Given an agent, β with a choice K and a proposition A if we observe instances of $\beta/K \rightsquigarrow A$ and instances of $\beta/K \not\rightsquigarrow A$ we infer that β/K may lead to A and write this as $\beta/K \Diamond\!\!\rightsquigarrow A$.

In addition to these two operators we extend standard modal logic by the introduction of an *other-agent* extension. In the standard, single agent, versions of modal rules and axioms there are a number of sentences that may admit multiple agents. The standard modal axiom C, for example, using $\rightsquigarrow$ in place of stit, may be written as equation 8 and multi agent extensions may be considered.

$$C. \quad [\alpha \rightsquigarrow: A] \wedge [\alpha \rightsquigarrow: B] \supset [\alpha \rightsquigarrow: A \wedge B] \qquad (8)$$

The standard single agent / multiple proposition statement becomes a multiple agent / single proposition statement. Casting C in this mould replacing the propositions, A and B with agents α and β and having those agents act sequentially on a single proposition gives C_{agent} which is written as equation 9.

$$C_{agent}. \quad [\alpha \rightsquigarrow: A] \wedge [\beta \rightsquigarrow: A] \supset [\alpha; \beta \rightsquigarrow: A] \qquad (9)$$

These agent extensions must be considered both as parallel cases, outlined above, and as serial cases. C_{agent}, in serial form states that $[\alpha \rightsquigarrow: A] \wedge [\beta \rightsquigarrow: A] \supset [\alpha; \beta \rightsquigarrow: A]$. Where C_{agent} holds it appears to indicate that α and β*'s* actions are not mutually exclusive.

Because the theory is based on observed evidence and is being investigated in a noisy setting, standard modal logic rules and axioms may not be falsified as cleanly as with strict stit. Noise may cause counter evidence to appear and this would cause a strict stit reading to fail. We examine two cases, the convergence axiom C and the rule of equivalence RE, chosen from a larger set of modal rules and axioms, as examples for a partial characterisation. These demonstrate the other agent extension in two very different settings. Taking C and C_{agent} as examples agents may generate the data of Tables II and III. Note that in the following tables N represents negative evidence as described in observation 3. This is represented by a tick because, as indicated in Section III-A we require only that we observe at least one instance of negative evidence, this is not tallied as we do for positive evidence, P, and counter evidence, C.

A. *Influence and the convergence axiom, C*

If an agent were to see to it that A holds and that B holds at an instant and its choice of action is the coincidental result of its having two independently reasoned goals then it does so without intending to see to it that A and B hold jointly. Neglecting the agent's intent, however, it would be difficult to deny that the agent does see to it that A and B do hold jointly and that the principle stated in equation 8 is supported. A noteworthy point here is that operators are constrained to an agent and choice pair, to say that $\alpha/K \rightsquigarrow A$ and $\alpha/K \rightsquigarrow B$ implies is that it is the same choice, K, that brings about both A and B. Assuming that equation 8 holds an attempt to generate counter examples is illustrated in Table II.

Table II
EVIDENCE SUPPORTING C

Hypothesis	Evidence P	Evidence N	Evidence C	P:C ratio	Conclusion
(IIa): $\rightsquigarrow$ example					
$\alpha/K \rightsquigarrow A$	n	✓	0	∞	$\alpha/K \rightsquigarrow A$
$\alpha/K \rightsquigarrow B$	n	✓	0	∞	$\alpha/K \rightsquigarrow B$
$\alpha/K \rightsquigarrow A \wedge B$	n	✓	0	∞	$\alpha/K \rightsquigarrow A \wedge B$
(IIb): $\Diamond\!\!\rightsquigarrow$ example					
$\alpha/K \rightsquigarrow A$	n	✓	p	n/p	$\alpha/K \Diamond\!\!\rightsquigarrow A$
$\alpha/K \rightsquigarrow B$	q	✓	r	q/r	$\alpha/K \Diamond\!\!\rightsquigarrow B$
$\alpha/K \rightsquigarrow A \wedge B$	s	✓	t	s/t	$\alpha/K \Diamond\!\!\rightsquigarrow A \wedge B$

In the noiseless case of Table IIa it can be seen that the hypotheses match the conclusion because of the lack of noise induced counter evidence, that is $C = 0$. It the noisy case of IIb, where $C > 0$, we apply the may lead to operator to the conclusion.

B. *The convergence axiom with agent extension, C_{agent}*

C_{agent} is the result of applying the other agent extension, as outlined above, to the standard C axiom and represents scenarios where a number of agents with potentially similar abilities act either in parallel or in series. This gives two versions, equation 10 says that if α can see to it that A and if β can see to it that A then α and β acting simultaneously can see to it that A holds. Equation 11 says that if α can see to it that A and if β can see to it that A then α and β acting serially can see to it that A holds.

$$C_{\|agent}.\ [\alpha \rightsquigarrow: A] \| [\beta \rightsquigarrow: A] \rightarrow [\alpha \| \beta \rightsquigarrow: A] \tag{10}$$

$$C_{;agent}.\ [\alpha \rightsquigarrow: A] \wedge [\beta \rightsquigarrow: A] \rightarrow [\alpha; \beta \rightsquigarrow: A] \tag{11}$$

$C_{;agent}$ is a potentially dangerous property. Consider agent actions that are mutually exclusive, for example, A may represent "pick up an indivisible token". If two agents, α and β have an action that may bring about A. Because coaching agent observations are limited to changes in the environment we must be careful that when both agents act so as to bring about A and although A may result α*'s* influence and β*'s* influence are mutually exclusive. The potential danger here is that a coach may identify this as a joint action when only one agent was responsible. Considering the parallel action case first, if α and β simultaneously act so as to bring about A then A will hold but it will do so as a result of α*'s* action or β*'s* action and not as a result of both actions. It is assumed that both agents have different abilities, that is that they execute different choices but these choices are functionally equivalent as far as A is concerned. If, say, $\alpha/K \rightsquigarrow A$ and $\beta/L \rightsquigarrow A$ and $\neg A$ holds then when agents act simultaneously both will perceive that A holds after their action. The data of Table III illustrates potential observed evidence for C_{agent}. Even in a single cell world where agents are always collocated the number of instances of $\alpha/K \rightsquigarrow A$ and $\beta/L \rightsquigarrow A$ will be greater than those of $\{\alpha/K \| \beta/L\} \rightsquigarrow A$. Assuming an even distribution of noise the *P:C* ratios n/p and q/r will be greater then s/t and ranking these as per the discussion in Section IV makes the single agent hypotheses appear to be more influential than the parallel action hypothesis. Returning to agents, there are two hypotheses for each – a single agent hypothesis which indicates an ability to bring about A and an other agent hypothesis indicating the same. Returning to the notion of gateways between domains of influence illustrated in Figure 4 it can be seen that the single agent hypothesis is contained in the single agent influence domain which is, by extension, contained in the two agent influence domain. After ranking,

Table III
EVIDENCE AND PARALLEL C_{agent}

Hypothesis	Evidence P	Evidence N	Evidence C	P:C ratio	Conclusion
$\alpha/K \rightsquigarrow A$	n	✓	p	n/p	$\alpha/K \Diamond\!\!\rightsquigarrow A$
$\beta/L \rightsquigarrow A$	q	✓	r	q/r	$\beta/L \Diamond\!\!\rightsquigarrow A$
$\{\alpha/K \| \beta/L\} \rightsquigarrow A$	s	✓	t	s/t	$\{\alpha/K \| \beta/L\} \Diamond\!\!\rightsquigarrow A$

the simpler single agent hypotheses are seen to carry more influence and offer a better account of α and β's individual ability to influence A than the two agent hypothesis.

If α and β operate sequentially then coaching agents will see much more evidence of influence for single agent action than for serial action. Given $\alpha/K \rightsquigarrow A$ immediately followed by $\beta/L \rightsquigarrow A$ the latter action will, except when noise intervenes, show no influence as A already holds and no change will be evident following β/L.

Whilst C_{agent} is not strictly falsified it may be seen that the foundations of the influence theory – observations of agent behaviour – do not lead to circumstances where C_{agent} may be considered as valid.

C. Rule of equivalence RE

Chellas [24] lists the rule of equivalence as:

$$RE. \ \frac{A \leftrightarrow B}{\Box A \leftrightarrow \Box B} \tag{12}$$

Casting this into influence based view gives:

$$RE. \ \frac{A \leftrightarrow B}{[\alpha \rightsquigarrow: A] \leftrightarrow [\alpha \rightsquigarrow: B]} \tag{13}$$

RE says that equivalent propositions are equally necessary. This relates to several aspects of agent behaviour in a coached environment. We must be careful with the notion of equivalence. If two propositions are absolutely equivalent, that is to say that they are the same but simply carry different labels or names, then one would expect to see evidence tallies matching exactly even in a noisy multi agent environment. If, however, the equivalence is both propositions are the result of the same choice then evidence tallies may not match exactly. In this case a coach would observe evidence supporting two hypotheses, one that a given action leads to A and another that the same action leads to B.

Let us assume two hypotheses, one that $[\alpha \rightsquigarrow: A]$ and one that $\neg[\alpha \rightsquigarrow: B]$ for α/K. Let us also assume that α/K brings about A and brings about B

Table IV
EVIDENCE SUPPORTING RE

Hypothesis	Evidence P	Evidence N	Evidence C	P:C ratio	Conclusion
(IVa): $\rightsquigarrow$ example					
$\alpha/K \rightsquigarrow A$	n	✓	0	∞	$\alpha/K \rightsquigarrow A$
$\neg\alpha/K \rightsquigarrow B$	p	✓	q	p/n	$\neg(\neg\alpha/K \rightsquigarrow B)$
$\alpha/K \rightsquigarrow B$	n	✓	0	∞	$\alpha/K \rightsquigarrow B$
(IVb): $\Diamond\!\!\rightsquigarrow$ example					
$\alpha/K \rightsquigarrow A$	n	✓	p	n/p	$\alpha/K \Diamond\!\!\rightsquigarrow A$
$\neg\alpha/K \rightsquigarrow B$	q	✓	r	q/r	$\neg(\neg\alpha/K \Diamond\!\!\rightsquigarrow B)$
$\alpha/K \rightsquigarrow B$	s	✓	t	s/t	$\alpha/K \Diamond\!\!\rightsquigarrow B$

In the noiseless example of Table IVa we see that the equivalence of A and B is reflected in the positive, negative and counter evidence tallies. Assuming that the $\alpha/K \rightsquigarrow A$ hypothesis holds, as in the statement above, and that A and B are equivalent propositions then there will be counter evidence for the $\neg\alpha/K \rightsquigarrow B$ hypothesis. The conclusion for the negative hypothesis, $\neg\alpha/K \rightsquigarrow B$, is that it does not hold.

The counter evidence that negates the $\neg\alpha/K \rightsquigarrow B$ hypothesis is evidence of a noisy environment where we consider a may lead to result. Here the P:C ratio for $\neg\alpha/K \rightsquigarrow B$ will be significantly smaller than that for both $\alpha/K \rightsquigarrow A$ and $\alpha/K \rightsquigarrow B$. Note that n/p and s/t are not necessarily equal, other agents may play a part in generating observed evidence.

RE considers equivalent propositions from the point of view of a single actor as the agent of change. In a multi agent environment the ability to extend a hypothesis across groups of agents, agents which are members of some equivalence class, will allow a coach to develop behaviours applicable to a greater number of actor agents.

D. The rule of equivalence with agent extension RE_{agent}

Before considering the agent extension to RE we address the question of *agent equivalence*, outlined above, which allows a coach to infer that one agent's ability to bring about A is transferable to other agents. Agent ability is driven by agent choice and we call similarly capable agents *choice class equivalent*. Given two agents, α and β, with choice sets $Choice_\alpha$ and $Choice_\beta$ respectively we say that α and β are choice class equivalent for a choice K iff $K \in \{Choice_\alpha \cap Choice_\beta\}$. Agents belonging to the same agent class will be choice class equivalent by default and the definition above may be extended across agent classes where these classes share common agent choices. This refinement of agent class equivalence to choice class equivalence removes a degree of coarseness from coach reasoning.

RE_{agent}, equation 14, maps RE's claim of the equivalence of propositions onto the domain of agents and allows the coaching operation to work with equivalent agent classes and to substitute equivalent actions in synthesised behaviours. The coaching operation treats agents as abstract entities, they are simply a collection of choices. It may be that different coaching agents use different representations of the same agent class, these representations may simply be different orderings of agent choices so that α/K and β/L amount to the same choice. At an agent class level RE_{agent} allows the coaching operation to aggregate data from several coaching agents each of which may have a different ordering on agent choice sets. Two agents from different classes may have equivalent choices, these may be the same choice contained in different agent choice sets or they may be different choices that lead to the same result. In such cases the coaching operation may substitute agent classes in a synthesised behaviour and this means that it may be possible for more sophisticated coaching agents to be able to optimise synthesised behaviours for agent populations.

$$RE_{agent}. \quad \frac{\alpha/K \leftrightarrow \beta/L}{[\alpha \rightsquigarrow: A] \leftrightarrow [\beta \rightsquigarrow: A]} \tag{14}$$

This is an important rule for our system, if RE_{agent} fails then a coaching agent can not assume that a good behaviour exhibited by one agent of a particular choice class may be transferred successfully to another agent of that class.

Table V
EVIDENCE SUPPORTING RE_{agent}

Hypothesis	Evidence			P:C ratio	Conclusion
	P	N	C		
$\alpha/K \rightsquigarrow A$	n	✓	p	n/p	$\alpha/K\Diamond\!\!\rightsquigarrow A$
$\beta/L \rightsquigarrow A$	q	✓	r	q/r	$\beta/L\Diamond\!\!\rightsquigarrow A$

Given two agents, α and β, belonging to separate agent classes and two different actions, K and L for α and β respectively, a coach may see evidence such as that of Table V. Both α and β present evidence of being able to bring about A and in this case the coach will simply consider this an equivalent ability for each agent class and seed a behaviour for each class. Functionally both behaviours are equivalent, the coach is unable to see a difference and simply treats them as equivalent.

VI. CONCLUSION AND FURTHER WORK

The motivation for this work was to investigate how agents may influence each other and to apply this concept of influence to an analysis of nested other agent behaviour. The investigation grew from earlier attempts at characterising emergent behaviour in simple systems of reactive agents, Logie et al. [3], which proved to to be a difficult problem and led to the consideration of deontic logics. This work is a step towards addressing that problem with new and more suitable tools.

The database structure of Figure 8 followed investigations into data mining sets of coaching agent observations, Logie et al. [2] and the realisation that branching may be collapsed into a potentially closed structure. The data of Figure 8 and Table I were generated by coaching agents observing actor behaviour in a noisy environment and one where only tiny amounts of influential α class agent behaviour would be evident. Given the noise and small incidence of influential α class behaviour these are impressive results (but always with the proviso that they are from a simple system).

Our partial characterisation indicates that, in common with stit, influence supports modal operators. More importantly the characterisation indicates that our theory of influence may be extended into domains requiring complex sequences of agent actions. The failed characterisations are equally important, the other agent extension indicated that our notion of influence rejects cases where agent actions may be mutually exclusive. These results bode well for further investigation to build a more solid understanding of exactly how a coach may manipulate evidence of uncertain ability without generating unrealistic conclusions.

This work, the theory of influence, implementation in a simple test system and a partial characterisation, is the first step in developing tools to investigate adaptive and emergent behaviour in systems of simple agents.

Future work will involve further logical analysis of influence by way of *leads to* and *may lead to* operators. We inferred nested influence from coaching agent database structure in this simple case but are unsure of what other inferences may be safely made. Further investigation into the logical properties of these operators will indicate what inferences may and may not be safe and this will allow experiments with richer and more complex environments where nested influence is not so obvious. Halbwach [23] indicated that the use of null actions was standard in programming synchronous reactive systems. This is something that needs to be investigated and has many interesting threads, how may a coaching agent detect that synchronisation can improve an aggregate behaviour? When and how do agents use synchronising actions and do they need more refined percepts to, perhaps, detect other agents?

Application of our influence theory in complex problem domains are being explored, notably software maintenance where the influence of a developer on a type or class of problem may be observed by a third party allowing for a recommendation system that works by observing existing defect and update management systems.

REFERENCES

[1] R. Logie, J. Hall, and K. Waugh, "Agent influence and nested other-agent behaviour," in *Proc. Future Computing, Service Computation, Cognitive, Adaptive, Content, Patterns, 2009. COMPUTATIONWORLD '09. Computation World:*, November 2009, pp. 488 –493.

[2] R. Logie, J. G. Hall, and K. G. Waugh, "Towards mining for influence in a multi agent environment," in *IADIS European Conference on Data Mining 2008 (part of MCCSIS 2008)*, A. P. Abraham, Ed. IADIS, 2008, pp. 97–101.

[3] R. Logie, J. G. Hall, and K. G. Waugh, "Reactive food gathering," in *CLIMA VI*, ser. Lecture Notes in Computer Science, F. Toni and P. Torroni, Eds., vol. 3900. Springer, 2005, pp. 406–413.

[4] R. Logie, "A study of agent influence in nested agent interactions," Ph.D. dissertation, The Open University, June 2009.

[5] R. A. Brooks, *Cambrian intelligence: the early history of the new AI*. Cambridge, MA, USA: MIT Press, 1999.

[6] N. Jennings, K. Sycara, and M. Wooldridge, "A roadmap of agent research and development," *Autonomous Agents and Multi-Agent Systems*, vol. 1, no. 1, 1998, pp. 7–38.

[7] N. D. Belnap, Jr. and M. Perloff, "Seeing to it that: a canonical form for agentives," *Theoria*, vol. 54, 1988, pp. 175–199.

[8] B. F. Chellas, *The Logical Form of Imperatives*. Stanford, California: Perry Lane Press, 1969.

[9] J. F. Horty and N. D. Belnap, Jr., "The deliberative STIT: A study of action, omission, and obligation," *Journal of Philosophical Logic*, vol. 24, no. 6, 1995, pp. 583–644.

[10] A. N. Prior, *Past, Present and Future*. Oxford: Oxford University Press, 1967.

[11] R. H. Thomason, "Indeterminist time and truth-value gaps," *Theoria*, vol. 36, 1970, pp. 246–281.

[12] R. H. Thomason, "Combinations of tense and modality," in *Handbook of Philosophical Logic, Volume II: Extensions of Classical Logic*, D. Gabbay and F. Günthner, Eds. Dordrecht: D. Reidel Publishing Co., 1984, pp. 135–165.

[13] B. F. Chellas, "Time and modality in the logic of agency," *Studia Logica*, vol. 51, 1992, pp. 485–517.

[14] N. Belnap, M. Perloff, and M. Xu, *Facing the Future: Agents and Choices in Our Indeterminist World*. Oxford: Oxford University Press, 2001.

[15] J. Ferber and J.-P. Müller, "Influences and reaction: a model of situated multiagent systems," in *In: Second International Conference on Multi-agent Systems.* AAAI Press, 1996, pp. 72–79.

[16] F. Michel, "The IRM4S model: the influence/reaction principle for multiagent based simulation," in *AAMAS '07: Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*. New York, NY, USA: ACM Press, 2007, pp. 1–3.

[17] D. Weyns and T. Holvoet, "A model for situated multi-agent systems with regional synchronization," in *10th International Conference on Concurrent Engineering, Agent and Multi-Agent Systems, CE'03*. Springer-Verlag, 2003, pp. 177–188.

[18] O. Simonin, A. Lanoix, S. Colin, A. Scheuer, and F. Charpillet, "Generic expression in B of the influence/reaction model: Specifying and verifying situated multi-agent systems," INRIA, Lorraine, Tech. Rep. RR-6304, 2007.

[19] J. F. Horty, *Agency and deontic logic*. Oxford University Press, 2001.

[20] L. Åqvist and P. Mullock, *Causing Harm*, R. Posner and G. Megle, Eds. de Gruyter, 1989.

[21] R. Milner, *Communication and concurrency*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1989.

[22] S. O. Hansson, "Situationist deontic logic." *Journal of Philosophical Logic*, vol. 26, no. 4, 1997, pp. 423–448.

[23] N. Halbwachs, *Synchronous programming of reactive systems*. Kluwer Academic Publishers, 1993.

[24] B. F. Chellas, *Modal Logic, an introduction*. Cambridge University Press, 1980.