A multi-agent model of hierarchical decision dynamics

Paul Kinsler*
Department of Electronic & Electrical Engineering, University of Bath, Bath, BA2 7AY, United Kingdom

(Dated: Monday 29th April, 2024)

Decision making can be difficult when there are many actors (or agents) who may be coordinating or competing to achieve their various ideas of the optimum outcome. Here I present a simple decision making model with an explicitly hierarchical binary-tree structure, and evaluate how this might cooperate to take actions that match its various evaluations of the uncertain state of the world. Key features of agent behaviour are (a) the separation of its decision making process into three distinct steps: observation, judgement, and action; and (b) the evolution of coordination by the sharing of judgements.

I. INTRODUCTION

Decision making has always been a potentially complex problem, and arguably never more so when there are many competing decision types to be made, when they apply to different scopes and arenas, when outcomes may be uncertain, and when there are many actors – with different levels of authority – who may be coordinating or competing to achieve their various ideas of the optimum outcome.

There is of course existing research in this area (e.g. decision making as briefly summarized in [1], and the advantages of hierarchies in [2]), but the deliberately abstract model presented here was constructed as a thought experiment and entirely without reference to any existing literature. Notably, the results from the model presented here do not focus on problem-solving or learning, but on how inter-agent communication and preferences affect the behaviour or performance of the system as a whole, how (or if) the system might converge to a final state in a static world, and how well-matched that state is to agent preferences, agent or network metrics, or to the world itself. It could, nevertheless, be interesting to add in a feedback step enabling the agents to adapt their preferences to better match the world. However, going as far as a formal representation – or remodelling – in terms of neural networks [3] is not my intent, although the analogy is certainly an interesting one [1].

Here I attempt a simple model of a process, focussing mainly on structurally hierachical organizations, where different levels in the hierachy are subject to different disadvantages and different time scales. This is a simple agent based model, specialised here to a binary tree structure, and contains agents with a specific class of decision mechanisms. Nevertheless, within the decision mechanism used, there is considerable freedom to adjust parameters, and the decision function used could easily be made more complicated or sophisticated. Here, however, I restrict myself to nearly identical agents using the same decision parameters.

One key feature is that the "decision" process is split into three distinct steps: information gathering, judgement formation, and action. Notably, any agent's judgement about a best action is not necessarily the same as the action taken, since (e.g.) the preferred action might be altered – or even overridden – by the judgements of higher level agents. The other key feature is that agents share only their judgements, and not their observations about the world, or their actions.

This paper contains:

-: a tree model for the interactions of a multi agent system which is explicitly hierarchical, both in terms of connectivity and speed of action.

-: Each agent takes into account the state of the world it inhabits, and that of its superiors and subordinates and combines these ...

-: to create (a) a judgement about what should ideally be done, and (b) a potentially dissimilar action that it will take.

-: This model is analysed on the basis of some "reasonable" assumptions for parameter values, for a range of cases, and some results showing how a network behaves on contact with a world unmatched to its expectations are presented.

-: In particular, role and relevance of differing success metrics are discussed.

II. MODEL

Here our decision hierarchy (network) is a binary tree containing $N = 2^n$ agents. This means the main interaction structure is of a parent agent with its two child agents; although any agent might fulfill both parent and child roles. As noted above, for simplicity each agent i behaves in the same way, and shares its *judgement* J_i about the best course of *action* A_i both up and down the tree.

The top level – the "highest" level – is level zero, and the agent there is the "ur-parent", and has no parent itself. For convenience we give this ur-parent the index $i = 0$. The bottom level – lowest level – "leaf" agents have parents, but no children. All other agents have both parent agents and child agents, and an agent interacts only with its parent or its children. A small example tree network is depicted on fig. 1.

---

* https://orcid.org/0000-0001-5744-8146

[1] I am even tempted to imagine a society as a kind of gas of interacting agents who take inputs from other agents and their surroundings and then provide summarised "opinion" outputs to others thus acting as a kind of weakly structured neural network (or perhaps "neural mesh"), but the structure is not set by a designed input-layer to output-layer architecture, but merely biased by the association preferences of the constituent agents.

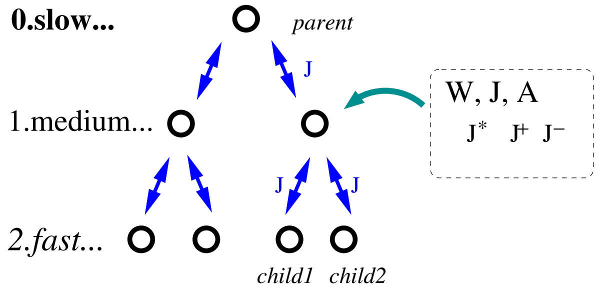

FIG. 1. A small three-level agent decision treee (agent-tree), indicating the typical agent state values as described in the main text, the sharing of judgements between agents, and the speed of operation of the different levels.

In what we present here we restrict all state, judgement, and action values to be real numbers.

Note that most of what is presented below generalises straightforwardly to other types of tree-like agent networks, whether binary or not, or symmetric or not. Even other non-tree networks might be modelled, although some adaptions will need to be made – i.e. imposition of a "level" property on agents, and reimagining the meaning of "parent" and "child".

A. Notation

We denote the judgement of an agent (with index) i at some time tick t_j to be $J_i(t_j)$, its action to be $A_i(t_j)$; these will be based in part on the agent's measurement $W_i(t_j)$ of the world state. If we wish to refer to any of the possible elements of agent state, we will use $Q_i(t_j)$; i.e. $Q \in \{J, A, W\}$. However, we will typically omit the time argument if it is not needed; typically this is because all quantities considered will have the *same* time argument.

Since it will be necessary to refer to the parents and children of an agent i, we denote the index of that parent as $p(i)$, and the indices of the children of i as $u(i), v(i)$. However, the top level agent (ur-parent) has no parent, and the bottom level agents (leaves) have no children. To address this we treat the ur-parent as its own parent, and the bottom level agents as their own children, so that in those specific cases, and only those cases we set $p(i) = i$, or set $u(i) = v(i) = i$.

Lastly, since the level of the agent, as determined by counting down the tree from the top (level 0), we denote this as $\lambda(i)$.

B. World

We call the environment that our agents operate the "world", and in most of the work presented here this is represented by a single real number $\mathcal{W}$, although an agent i's measurement W_i of that state may be inaccurate.

It is convenient to think of a state value of zero as a situation requiring no action ($\mathcal{W} = 0$), a judgement that no action is necessary ($J_i = 0$), or an agent taking no action ($A_i = 0$); and that values deviating from zero are somehow "more extreme" the larger the value. If $\mathcal{W} = J_i = A_i$ then we should imagine that the judgement J_i is exactly appropriate to the world situation $\mathcal{W}$, and that J_i's commensurate action A_i is likewise also somehow perfectly matched to the situation.

It is simplest to assume that actions taken by the agents do not affect the world, which means that any investigation focusses on how well some specific agent-tree will adapt to the world it inhabits. However this is not a requirement, so:

XW: *Optionally*, the actions of each agent can "hammer" on the world and change it. Each hammer blow (change) is proportional to that agent's $(A_i - \mathcal{W})$, divided by the number of agents, and multiplied by some prefactor ε.

C. Agent state

Each node (agent) in the tree has three items of local state. Here we consider the case where each element of state is simply a single real number, and if the values match they "agree" in some decision sense, and any mismatch indicates a proportional degree of disagreement. Specifically, for any agent i these state elements consist of three local ones:

S1.: its observation W_i of the world value $\mathcal{W}$,

S2.: its (most recent) judgement J_i of what to do, which ideally would match $\mathcal{W}$ exactly,

S3.: its (most recent) action A_i taken, which ideally would (also) match $\mathcal{W}$ exactly;

as well as three state elements from its neighbours:

S4.: the most recent judgement of its parent $J_i^* = J_{p(i)}$,

S5, S6.: the most recent judgements of its children $J_i^{\pm} = J_{u(i)}, J_{v(i)}$,

An agent's state is thus contained in a six-vector:

$$\boldsymbol{\omega}_i = \left(W_i, J_i, A_i, J_i^*, J_i^+, J_i^-\right) \tag{1}$$

Here we use the auxilliary definitions J_i^*, J_i^+, J_i^- because any agent i learns or gathers these at a particular time, so they remain fixed until again updated; whilst the parent or child agent may update their actual J values independently.

D. Agent algorithm

Agents follow a three-step process, where each step enables the agent i to update its state vector $\boldsymbol{\omega}_i$ in the sequence

T1.: measure the world ($\mathcal{W}$), to obtain an updated value for W_i, and collect a new values for its parental judgement $J_i^* = J_{p(i)}$ and child judgements $J_i^+ = J_{u(i)}$, $J_i^- = J_{v(i)}$, and update $\boldsymbol{\omega}_i$;

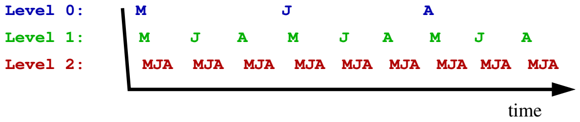

FIG. 2. Timing diagram: higher level agents act first, but lower level agents act more frequently. The step labelled M is the measurement step **T1**, that labelled J is the judgement step **T2**, and that labelled A is the action step **T3**.

T2.: use its current state $\boldsymbol{\omega}_i$ to decide on a new judgement J_i', and again update $\boldsymbol{\omega}_i$;

T3.: use its current state $\boldsymbol{\omega}_i$ to take its next action A_i', once again updating $\boldsymbol{\omega}_i$.

E. Hierarchy

The binary tree model used here is already hierarchical. However, in addition to this intrinsic behaviour we add two extra properties.

First, to incorporate the different timescales desired for (slow) high level agents and (faster) lower level agents, for every one step a parent agent takes, its children complete all three steps (T1,T2,T3); and in any single time tick parents act before children [2] The relative timings of these steps for agent of different levels are indicated in fig. 2. Although alternative timing schemes might be imagined, here we choose this one because of the interleaving of different types of steps taken at different levels.

Second, we allow measurements of the world value $\mathcal{W}$ to have an accuracy that varies with agent level, as controlled by a parameter ψ, a feature which will be addressed in detail below.

F. Measurements, Judgements, and Actions

1. T1: Measurements

A *measurement* W_i of the world by an agent i returns the current (global) value $\mathcal{W}$, plus some random noise; i.e.

$$W_i = \mathcal{W} + \eta \psi^{\lambda(i)} \xi, \tag{2}$$

where ξ is a random variable with zero average $<\xi>= 0$.

Here we imagine that in most cases, lower level agents see proportionately more noise, reflecting (perhaps) the larger variation in their local conditions; as opposed to the coarse-grained averaging expected of higher level decision makers. Thus the amplitude of the added noise varies from one level to the next lower according to a proportionality constant ψ; e.g. at three levels down the relative noise multiplier is ψ^3. This scaling is chosen to make parent-child comparisons of noise contributions self similar between different levels. Note that *if* we preferred low level agents to (e.g.) "see more clearly", we could set $\psi < 1$, adapt η appropriately, and have low level agents experience small measurement noise and high level agents seeing larger noise.

For any agent i, collection of parental judgements $J^* = J_{p(i)}$ and child judgements $J_i^+ = J_{u(i)}$, $J_i^- = J_{v(i)}$ is taken to be perfect and without error.

2. T2: Judgements

A *judgement* J_i is a weighted sum over an agent i's six state elements. There are thus six possible judgement parameters. It is *arguably reasonable* that any agent's judgement is most strongly influenced by its own W_i, but the model below requires no such restriction.

Here we specialise to a case where all the agents to have the same judgement weights, which will be defined to match the "arguably reasonable" idea above, so that

$$\boldsymbol{\sigma} = (1-3\theta, 0, 0, \theta, \theta, \theta), \tag{3}$$

where our default value will be $\theta = 1/10$. Note that this choice of $\boldsymbol{\sigma}$ ignores an agents past judgements and actions, and is thus *memoryless*; and since it sums to unity can be coonsidered *conservative*.

Here we will restrict ourselves to a simple linear computation of a judgement

$$J_i' = \boldsymbol{\omega}_i \cdot \boldsymbol{\sigma}_i, \tag{4}$$

which means that we then must update $\boldsymbol{\omega}_i$ to incorporate the new J_i value by replacing the old.

3. T3: Actions

An *action* A is also a weighted sum of an agent i's state. There are thus six possible action parameters. It is *arguably reasonable* that any agent's actions are most strongly influenced by its parent's judgement, but the model below requires no such restriction.

Here we specialise to a case where all the agents to have the same action weights, which will be defined to match the "arguably reasonable" actions above, so that

$$\boldsymbol{\alpha} = (0, \phi, 0, 1-\phi, 0, 0), \tag{5}$$

where our default value will be $\phi = 2/10$.

Note that this choice of $\boldsymbol{\alpha}$ ignores an agents past judgements and actions, and is thus *memoryless*; and since it sums to unity can be considered *conservative*.

Here we will restrict ourselves to a simple linear computation of the chosen action

$$A_i' = \boldsymbol{\omega}_i \cdot \boldsymbol{\alpha} \tag{6}$$

[2] Arguably, this should be different, with e.g. the three steps being treated as a unitary whole, and rather than the 1:3 ratio defined by the three steps, be some thing else, perhaps 1:2.

which means that we then must update $\boldsymbol{\omega}_i$ to incorporate the new A_i value by replacing the old.

G. Missing features

A significant omission in this deliberately simple model is any sense of logistics or action constraint; which is problematic if the world state is modified ("hammered") by agent action. Since we assume actions A are somehow more extreme if they have larger values, placing some limit on the rate of A expended would seem reasonable; and also mollify any tendency of a runaway to ever larger $\mathcal{W}$ by strategies that promote – by accident or design – increasingly aggressive agents.

Of course, several suitable modifications suggest theselves, but all require extra complexity in a model that aims at simplicity, but nevertheless already has many parameters. Therefore we defer treatment of this to later work.

Nevertheless, perhaps the simplest logistics constraint would be to add one extra element of agent state, namely fuel F_i. This could be incremented at every time tick, perhaps only up to some maximum F_{MAX}. Then an agent i could only take some action A_i if $F_i > A_i$, after which the fuel F_i would be decreased by A_i. Note that such extra detail should be an *internal* feature of some more sophisticated action determination (II F 3), rather than an explict additional part of the algorithm proposed here.

III. INITIAL CONDITIONS, PARAMETERS, METRICS

1. Initial conditions

Many possible initial conditions could be considered, but here we start all agents with W, J, A all zero, i.e. inactive. The world starts at some non-zero finite value (eg $\mathcal{W} = 3$), so that the tree of agents has some discrepancy to adapt to.

2. Parameters

Here we choose our default measurement noise proportionality to be $\psi = \sqrt{2}$, and the prefactor to be $\eta = 10^{-3}$. Here our random variable ξ is drawn from a rectangular distribution spanning the range $[-1,+1]$.

If present (i.e. not zero), we choose the "hammer" parameter to be "small" (e.g. 2×10^{-3}), and scaled as $1/N$. (remember $N = 2^n$).

The standard judgement parameters $\boldsymbol{\sigma}(\theta)$ and action parameters $\boldsymbol{\alpha}(\phi)$ were given before in (3) and (5) as part of the discusssion when they were introduced.

3. Metrics

Measures of success are here related to the summed squares of differences between agent actions A and some chosen state Q. Thus for agents indexed by i, we have success measure X based on Q as

$$X_Q = \sum_{i=0}^{n} (A_i - Q_i)^2 . \tag{7}$$

Keeping in mind that the most success matches the case when the measure sums to zero, some possible measures that might be considered are:

D0: naive success is when each agent's actions A_i match its observations, i.e. when $Q_i = W_i$,

D1: absolute success is when actions A_i match the world value, i.e. when $Q_i = \mathcal{W}$ – although of course no agent or set of agents could reliably calculate this,

D2: perceived success is when actions match judgements, i.e. when $Q_i = J_i$,

D3: bootlicker's success is when actions match an agent's parent's judgement, i.e. when $Q_i = J_i^*$,

D4: authoritarian success is when actions match the ur-parent's judgement J_0

D5: democratic success is when actions match the agent-tree's average judgement $Q = \bar{J} = \sum_i J_i$.

We can immediately see that D2 and D3 might be trivially satisfied by appropriate choice of $\boldsymbol{\sigma}$ or $\boldsymbol{\alpha}$. However, such choices will be unlikely to assist if absolute success is also under consideration.

Other metrics also suggest themselves. An agent tree in a malleable world may have a target value of $\mathcal{W}$ it wishes to achieve, or many other possibilities.

An important point raised by such a selection of success criteria is how to balance the competing demands of several at once; bearing in mind that the control parameters are $\boldsymbol{\sigma}$ and $\boldsymbol{\alpha}$, and the results are combined (or mangled) by the hierarchical judgement sharing and the different timescales.

IV. OUTCOMES

Here we assume *arguably reasonable* agents, as discussed in the preceeding section: those who base their judgements mostly on their observations, but whose actions tend to follow orders from their superior, as motivated by the preceeding discussion. This behaviour is specified by our default parameters.

In the following figures, the plotted quantities and axis scales are designed to normalise out the gross effects of the tree size (i.e. the number of agents N in the tree).

A. A clear and invariant World

The first case we look at will be for perfect vision (no observation noise) and an unchanging world state $\mathcal{W}$ ("no hammer").

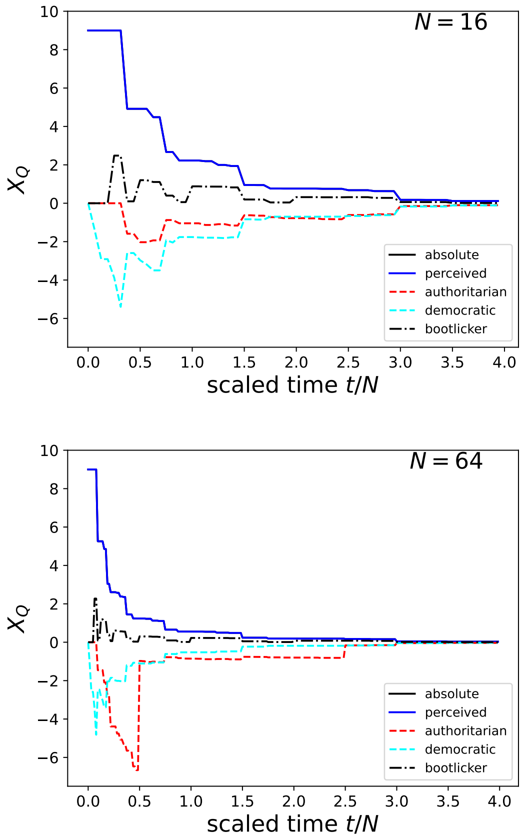

FIG. 3. Noiseless nohammer: results for a small four-level tree with $N = 16$ compared to a larger six-level tree with $N = 64$.

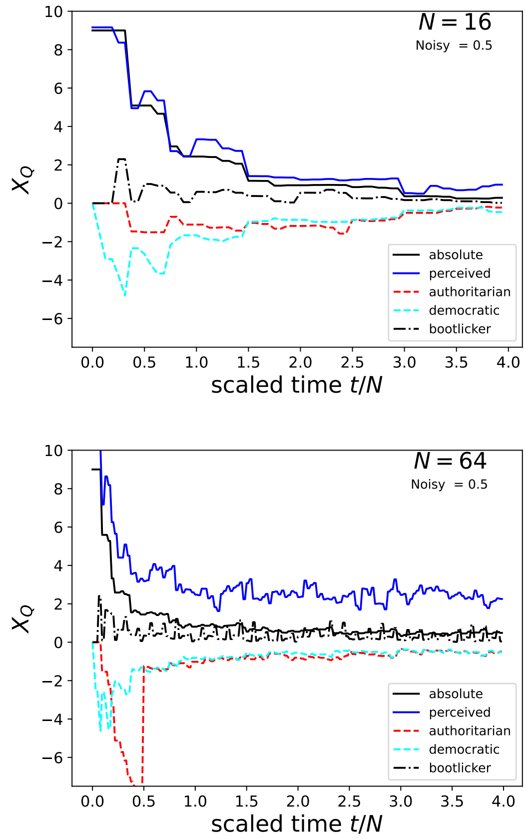

FIG. 4. Noisy nohammer: results for a small four-level tree with $N = 16$ compared to a larger six-level tree with $N = 64$. The size of the observation noise parameter is indicated on the panels.

As a consequence the behaviour of the system is relatively striaghtforward – on fig. 3 we see an essentially direct convergence to high success for all metrics, with $X_\mathcal{W}$ and X_J being identical. The structure (jumps and transitions) seen on the results are deterministic, and a result of distinct transitions caused by the algorithm and its interleaved timing.

B. A foggy but invariant World

This second case introduces observation noise, but still enforces an unchanging world state $\mathcal{W}$.

In this case, as shown on fig. 4, we again we see a convergence to a state with excellent success metrics (i.e. zero valued). However we see an evolution that whilst similar to that on fig. 3, is affected by noise, and exhibits a moderate discrepancy between the network's absolute success $X_\mathcal{W}$ and its perceived success X_J.

Since the perceived success does not converge to the absolute success; it remains worse (and by by a non-negligible margin) than all the other measures shown. This means that the agents will – on average – remain more dissatisfied with their performance than they should be.

C. A clear but malleable World

In the third case, we incorporate only a feedback (hammering) of agent actions on the world state. Because I have chosen parameters where the affect of the world is small, here we see that fig. 6 appears almost identical to fig. 3; but there is a key difference.

Close examination, and as recorded by the ΔWorld value on the panels, the world state value $\mathcal{W}$ is slowly but inexorably increasing with time. This is indicated by the slight gradient on the near-horizontal dotted line. This runaway escalation in decision making behaviour can be reduced or enhanced by altering the simulation parameters; here they are tuned to keep the increase gradual.

Note that trying to moderate the escalation by using non-conservative judgment or action weights (i.e. where $|\boldsymbol{\sigma}|$ or $|\boldsymbol{\alpha}|$ are scaled to less than unity) is surprisingly ineffective. This is likely because any mismatch between W, J, and A will always tend to make things worse, but it deserves closer investigation.

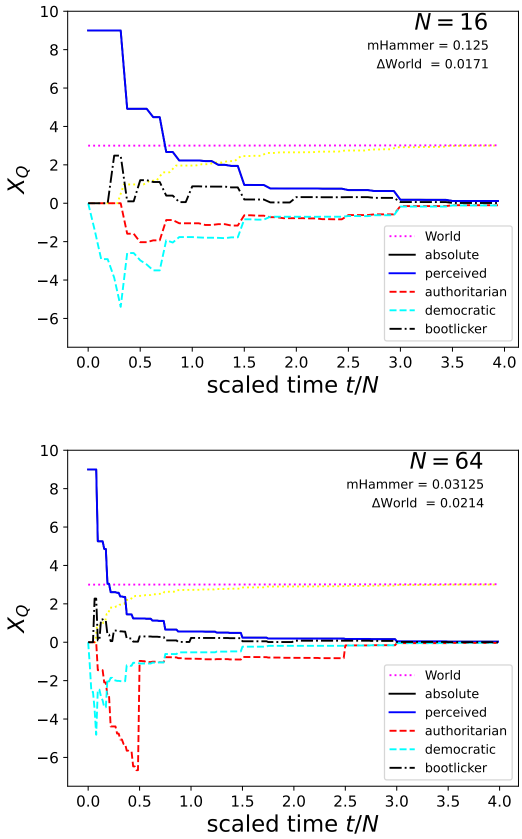

FIG. 5. Noiseless hammer: results for a small four-level tree with $N = 16$ compared to a larger six-level tree with $N = 64$. The size of hammer parameter is indicated on the panels. The near-horizontal dotted indicates the gradually increasing value of the world state $\mathcal{W}$; the numerical size of the change is indicated with the label ΔWorld.

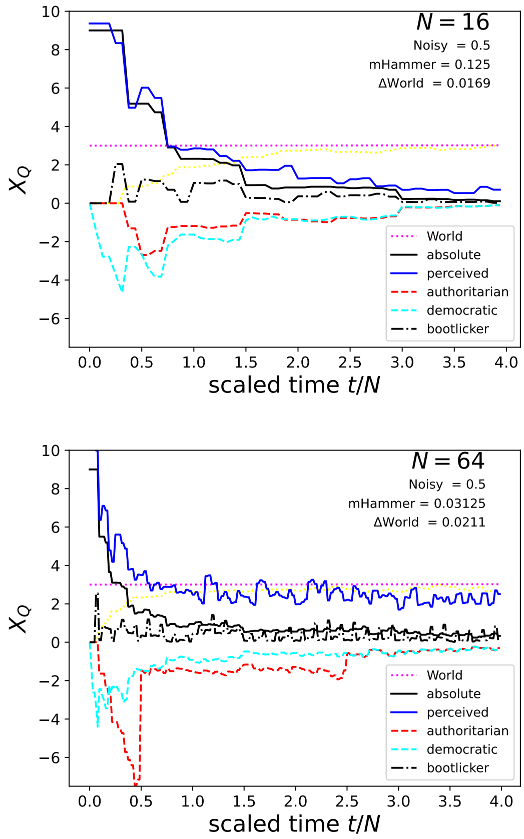

FIG. 6. Noisy hammer: results for a small four-level tree with $N = 16$ compared to a larger six-level tree with $N = 64$. The size of the observation noise and hammer parameters are indicated on the panels. The near-horizontal dotted indicates the gradually increasing value of the world state $\mathcal{W}$; the numerical size of the change is indicated with the label ΔWorld.

D. A foggy and malleable World

In the fourth and final case we incorporate both observation noise and feedback (hammering) of agent actions on the world state.

On first sight of fig. 6, these are perhaps not dissimilar to a noisier version of the “Noisy Nohammer” results of IV B, but there is the same key difference as for the noiseless hammer results presented in Sec. IV C.

First, however, we see again that the network’s perceived success does not converge to the absolute success; it remains worse than all the other measures shown, although the margin is unchanged by the new “hammer” feature in the model. Again, the agents remain more dissatisfied with their performance than they should be.

Second, again we see world state value $\mathcal{W}$ slowly but inexorably increasing with time, as indicated by the positive valued ΔWorld number.

V. FUTURE POSSIBILITIES

There is obviously a great deal of scope – even within the model as presented here – to investigate a wider range of parameters. Less conservatively, there are also many extensions, such as adaptive agents, allowing diversity of agent properties, or more complicated and even dynamic judgement algorithms.

Nevertheless, between these two extremes some interesting possibilities suggest themselves; and would be interesting to pursue once a classification of the judgement and action parameter $\boldsymbol{\omega}, \boldsymbol{\alpha}$ state space has been made. Some possibilities are:

Changing world: Set the world $\mathcal{W}$ to have an oscillation, and investigate the interplay of the decision time scales against the world’s timescales. The more unsympathetic of us might instead test their network against a more complicated (or even chaotic) variation in $\mathcal{W}$.

Competing networks: Have a pair of networks side by side, and have the agents hammer on their counterpart – so we have that network I’s agent action A values being network II's agents world W values, and vice versa. You might also try putting the networks foot to foot, so that only the lowest level agents see each other.

Strategic vs tactical: Have a world function where high level agents see different W values to low level ones (perhaps even with opposite sign). In such a circumstance, how might a network come to a sense of agreed action?

VI. SUMMARY

This hierarchical decision model is deliberately simple, but attempts to incorporate the essential features of hierarchy (multiple levels), different decision time-scales, the variation in perceptions, and inter-level communications.

Thus, in its favour, it incorporates those design features, whilst also having configurable (and generalisable) judgement and action steps. Further, since it is the behaviour of an agent within the network which is the core algorithmic feature, it might be generalised to different sorts of trees or other network configurations.

However, the model does have some limitations: notably, that the decisions themselves are handled in a very abstract way, and are all of similar type; there is no mechanism to make high-level (which we might want to consider "strategic") functionally different to the mid- or "tactical" low-level ones. Further, and in common with many abstract models, whilst one might add more features or improved algorithms, such added complexity will not obviously make the model more realistsic, but may only obscure interesting generic behaviours.

The selection of success metrics proposed here, and their explicit comparison, is key to the understanding of evaluating decision processes: the same decision (or decision process) can have outcomes that are regarded differently by different actors or agents – even those who are on the same side. Indeed, in several cases above we saw that the perceived success can quite easily be less than the *actual* success.

ACKNOWLEDGMENTS

The author would like to acknowledge the role of the Maths of EMA workshop run by PA Consulting in early 2023. One of the topics discussed was relevant to this field, and spurred me to construct this model, and to do this brief investigation.

[1] C. Gonzalez, P. Fakhari, and J. Busemeyer, Dynamic decision making: Learning processes and new research directions, Human Factors **59**, 713 (2017).

[2] N. Halevy, E. Y. Chou, and A. D. Galinsky, A functional model of hierarchy: Why, how, and when vertical differentiation enhances group performance, Organizational Psychology Review **1**, 32 (2011).

[3] D. Lin, S. Xue, D. Liu, M. Liang, and Y. Wang, Adaptive dynamic programming-based hierarchical decision-making of non-affine systems, Neural Networks **167**, 331 (2023).