Coordination of Many Agents in a Power System

Mohamadou Bella Bah*

Abstract

I present a model of negotiation between autonomous agents in the context of the U.S. electric grid. The model describes the interaction protocols employed by agents to reach agreements concerning price, contract duration and time of delivery. I go on to describe the strategies and tactics used to evaluate proposals and present counter proposals.

1. Introduction

Electricity is peculiar in that it needs to be consumed the moment that it is created.[1] If the grid cannot accommodate an additional injection of power at some location, a generative asset elsewhere is curtailed. Curtailment is either a strict loss for the generative asset (who profits from selling its output), or a loss for the utility, which may be forced, as in the case of renewables, to compensate the asset. Ostensibly, more favorable settlements could be reached were the suitable platform to exist: i.e. the generator could coordinate with storage to absorb its output; or with large but infrequent loads. Siloed business processes and a lack of integration of disparate brands and vintage of hardware and software means such settlements only emerge when the entities in question are the property of a single operator. This leaves grid operators with the daunting challenge of economically and reliably meeting the load demands of diverse independent customers within the physical limits of their equipment. And the physical limits and capital costs are onerous; hidden electrical coupling on the grid makes for nasty surprises. For instance, when a utility manipulates electricity demand independently of other grid controls, subsequent load changes make prior grid-regulation settings obsolete; compromising their ability to compensate for the adverse effects of the aforementioned manipulations. Potential impacts include voltage violations on the affected feeder, breaker trips and reductions in the effectiveness of corrective actions (read: manipulations) by as much as 15%! As you can imagine, the mismatch between the timescales of distribution grid controls, which operate within five minutes plus, and photovoltaics and inverter power electronics, sub seconds to milliseconds, exacerbates the issue. (Conventional feeder controls remain slow to adjust).

The Supervisory Control and Data Acquisition (SCADA) systems used to manage transmission switching, communications and monitor grid-state[2] generates *coarse data*: data that is mostly implicit information. It is an engineer's job to consolidate data sources and extract from them actionable insights. The manual analysis is tedious enough to warrant a strong desire for automated grid controls and data analytics. With that said, because real-time analysis is not offered on the distribution grid[3], operating staff make predictions about the grid's controllability (i.e. line switching, customer load management, etc.) based on past system studies into the estimated bus voltages/currents for a limited number of contingency scenarios.

The conventional notion of a unidirectional power system is quickly becoming incompatible with today's environment. Focuses and attitudes are shifting: Content with minimal outages[4], customers are now more concerned with power quality and the environmental impacts of their electricity service. The growing rise of 'prosumers', entities that interchangeably consume and produce power, has invited thought into the ways they might safely offer ancillary services to the grid. This includes the mechanisms governing the propagation of pricing information and the design of controls over loose and independent resources. (Prosumer/private owned photo-voltaic (PV) and DC-AC inverters can supply voltage and reactive power regulation services on the distribution feeder to preserve stability. The

*M. Bah is an Electrical Engineering and Computer Science student, Massachusetts Institute of Technology, Cambridge MA bellabah@mit.edu

[1] Storage relaxes this constraint somewhat.

[2] Refers to the knowledge of the voltage phasors at all buses in the system relative to a reference.

[3] Embarrassingly, utilities still rely on customers to let them know of substandard service or other interruptions.

[4] The grid was designed with keeping outages minimal. This was achieved through grid redundancies and ensuring permanent excess capacity.

challenge is in doing so irrespective of the magnitude, location and timing of the power injections). Markets at the distribution level are outclassed by the growing maturity of prosumer-permitting technologies like electric vehicles (EV), solar and comparably intermittent distributed energy resources (DERs).

DERs pose stability problems. As DERs like PV supplant incumbent synchronous generators, grid stability is negatively affected by the decline in overall system inertia; which otherwise suppresses the random fluctuations in line power flows and voltage disruptions on feeders. The interplay of unpredictable power flow physics and the network's dynamic topology can lead to outages or cascading failures on the grid.

Moreover, DERs, rather than being welcomed additions to the grid, are viewed by utilities as nuisances to be accomodated; treated as a negative load (non-dispatchable generator) under current policy, DERs are exempt from the same under-generatoion penalties [6] their conventional fossil-fueled cousins face. Solar is especially troublesome. Its reputation is marred by the infamous duck curve, a net load curve obtained from subtracting solar's unique production curve from the consumer load curve. The curve is characterized by steep tall ramps at either side of a deep valley that resembles a duck's side profile. Utilities prefer 'flat' curves wherein baseload demand may be serviced by cheap power plants running 24/7. Because familiar renewables cannot accurately forecast production[5], utilities, under renewable portfolio standards, keep dependable generative assets like coal and natural gas power plants on standby. This is not the optimal use of resources.

1.1. Physical Layout

It is not possible in modern grids to force electrons along predefined paths[6]. (Desirable because shortest-distance algorithms could be applied to minimize transmission losses and congestion; not to mention that it implies simpler power flow calculations. Ideally, physical laws take care of this, but the dynamic and interconnected nature of the electric grid foils nature's energy minimization efforts). Because it is impractical to modify the grid such that direct a-b paths can be created over which agent-a and agent-b transact[7], I settle for using the grid like an exchange. A simple protocol for transaction would be as follows: agents agree on the issues of the contract (i.e. price, volume, duration, etc.). Agent-a discharges the agreed amount into the grid as per the contract, meanwhile agent-b 'purchases' said amount from the grid. The monetary aspect of the transaction is settled the way you would expect (blockchain technology may be applicable here), with the 'grid'[8] possibly taking a commission. Because the grid can be unexpectedly fragile and unpredictable, it is unclear whether numerous but small transfers could be withstood without issue. (It is thought that the cumulative effect of small unusual fluctuations in a network could be the culprit of events like line failure[9].) Although more accurate, repeating this method of exchange with the grid as a middleman is cumbersome; therefore, when I talk of bilateral negotiations between agents, it is simpler to assume a super-wire connects the agents in question. (Super because it is bi-directional and has an unlimited energy transmission capacity.) Furthermore, I deliberately confine detailed physical discussions of the grid (this includes considerations of transmission, voltage and frequency constraints)) to this section so that it may not distract from the more algorithmic parts of the paper.

The normal grid looks like the following market-wise: It is composed of a day-ahead market (DA) and a real-time energy market that settles the imbalances between the DA's commitments and the actual output of its dispatched generators. The DA market is auctioned, wherein power producers submit quantity-price bids to the independent system operator (ISO). The ISO creates a supply curve and determines the equilibrium (read: market) price and quantity through the intersection of the supply curve and some predetermined demand curve. An energy management system (EMS) manages things like state estimation, network topology and voltage/frequency control. Grid operators normally ascribe to the grid five operating states: normal, alert, emergency, in extremis and restorative [10]. All grid states besides normal are technically bad and require corrective or *heroic* actions. Corrective action is undertaken actively by operators, or passively through

[5] Forecasting techniques exist to alleviate this issue, but none provides enough certainty to overcome the grid's caution. Some researchers are skeptical perfectly accurate and precise forecasting methods will ever exist.

[6] FACTS (flexible alternating current transmission systems) is an interesting development and fits well with the proposed models to come.

[7] I have some reservations about this, but I bow to the general opinion.

[8] I refer to the grid vaguely here. It is up to stakeholders to decide who receives the commission. I would lean towards transmission/distribution (T&D) as having a more defendable claim. Commissions here would go towards servicing the standing costs of the physical T&D infrastructure.

[9] local behavior emerges from the local interactions of the system's constituents.

automated feedback systems on the grid.
I now go on to present methods of control intended to make you attuned to later parts of the paper that extol the merits of many-agent coordination.

1.2. Transient Stability Control

This is the ability of the power system to maintain synchronism (one measure of grid-stability) in spite of large transient disturbances like bus faults or sporadic weather events.

Load changes in a grid, the result of topology changes (i.e. disconnecting loads, opening circuit breakers, line outages, the connection of electric vehicles, etc.), can cause rotating generators to accelerate or decelerate to match power imbalances, in turn compromising the system's frequency stability. More precisely, within AC systems, generators are electrically coupled together with outputs roughly in phase. (A group of such generators form a *synchronous area*). When local electric load increases, a feedback device called a governor offsets the generator's slowdown through the adjustment of a steam flow control valve. If the governor were faulty, it would fall upon neighboring generators to pick up the flagging generator's slack – thus bringing down the system's frequency. When load changes are particularly severe, frequency stability is degraded, and cascading outages can ensue. In extreme cases, generators spin rapidly to the point of mechanical failure.[10] The adjustment of a generator's fuel intake or output is currently the principal method of grid control – automatic generation control (AGC).

** Insert Figure **

In the two-machine system above, if the line connecting bus-1 to bus-2 were disconnected, the load-generation imbalance would damage the generators. If line reclosure was delayed, what could a system operator do in the meantime? One method is an insertion of a 300 MW brake resistor R at bus-1 and the removal of 300 MW of load at bus-2. This achieves the desired *effect* of power balance at each unit. Once the line is closed, the brake resistor is to be disconnected and load restored at bus-2. The difficulty lies in quickly shedding x amount of load reliably – fast-load-shed. Contracts and financial incentives that permit a utility to control the electricity consumption profile of human cosigners in return for reduced rates exist, but may be unreliable in that it is left to the cosigner to respond appropriately. Ask yourself if you would unfailingly obey a utility's call to disconnect your electric vehicle (EV). Application success in human-in-the-loop systems depend on a person's acceptance of control actions. Prospect theory, a descriptive model, suggests that human decisions are the result of heuristic evaluations of losses and gains[11]. Ideally, a balance between human involvement and automated controllers can be struck. It is worth thinking of how one could develop systems that respond to such scenarios automatically through the aggregation of privately owned grid devices –

1.3. Electricity Hopping

Here I propose something that is most likely unfeasible but is interesting nonetheless for its emphasis on decentralized/distributed operations.
Taking the section above literally, the limitations are obvious: the option of attaching some brake resistor is wasteful and finding some flexible local load to absorb 300 MW for an unspecified duration is difficult, unless of course there is perfect coordination. To this end, I propose the aggregation of residential storage for *electricity hopping*, a method inspired by the way in which data sent across the internet is split into packets and sent node-by-node. Similarly, one could split x MW of electricity and hop it via storage until it either reaches an area whose demand exceeds supply, or by some luck find its way back to its place of origin. (Do not take this too literally, ICT systems and the electric grid are different beasts.) The point of this proposal is to afford a utility time. The load shifting practiced today can be described as stalling or load displacement – typically until demand picks up again[12]. Residential storage here acts as a buffer to minimize transmission congestion, as well as avoiding other grid-stability issues. Most important however, is that it artificially introduces time delays. The way I imagine it, a utility should avoid transferring large quantities of energy at the same time via an ill-suited distribution network wherever they can. An intelligent storage device would of course have to correctly identify when it can store energy without compromising user comfort, and when to release that energy back into the grid. (Although broadcasting the time of release could be the utility's responsibility). A storage device would be compensated for undergoing life-shortening charge and discharge cycles. In a way, this concept is not so much

[10] For more information, search up the *Aurora Generator Test*.

[11] Studies demonstrate that people perceive losses differently to gains, tending to be asymmetrically loss averse. Not to mention we are paradoxically poor at fully appreciating non-linearities into the future despite our daily perceptions being logarithmic.

[12] As of now, daily energy demands are rather predictable.

’hopping’ as much as ’buffering’. The effectiveness of said proposal relies on timing among other things but once again brings effective-coordination to the fore.

1.4. Neighborliness

This proposal will help introduce multi-agent systems via agents that possess storage units.
Suppose there are n agents of varying states of charge (SOC) disconnected from the main grid but attached to a shared electrical bus. It may help to imagine the agents being EVs. Energy is transferred between agents via the shared bus and energy transmission within it is lossless, although energy in the system overall is not conserved, i.e. an agent may bring in energy from elsewhere and add it to the system, similarly, agents expend energy by doing work outside the system. Each agent has some historical record H of its past use from which it derives the *target level of charge* (TLC), a result also informed by its desired state of charge, battery’s state of health[13], and the current time. For now, one could imagine the TLC being related to urgent-use requirements, if an electric vehicle at $x\%$ SOC anticipates travel in an hour, it would understandably have a higher TLC than another agent that intends to leave in five hours with an SOC also of $x\%$. Agents also know the rated capacity C of their battery. Rather naively, we may say that whenever $soc_m < tlc_m$, general agent-m acts greedily and probes nearby agents whose state of charge exceeds their personal target level. With that, we build a set S composed of (agent, difference between state of charge energy and target-energy) tuples described below:

$$S = \{(agent_k, soce_k - tlce_k) | soce_k > tlce_k, k \neq m\} \quad (1)$$

For all eligible agents, order them from max-min by their energy difference $soce - tlce$ which we denote Δ. We want to find $X = \{x_1, x_2, \ldots, x_{|S|}\}$ that minimizes the sum of the ratios described below:

$$\text{find } X = Min_{x_i} \sum_{i=1, i\neq m}^{|S|} \frac{x_i}{\Delta_i} \text{ such that } Sum(X) = \Delta_m \quad (2)$$

$|\cdot|$ denotes the cardinality. The equation above refers to the little amounts taken from each eligible agent towards fulfilling agent-m’s deficit, i.e. x_1 refers to the amount of charge taken from agent-1. We use such a partitioning method to make it ’fair’. This may be difficult to implement in practice because as $|S|$ increases, it becomes increasingly complicated to discover X because there exists many parameters x which may be varied.

[13] State of charge refers to the battery’s charge relative to its current capacity, while state of health is a measure for the battery’s deterioration in its capacity since its manufacturing.

2. Multi-Agent Systems

I’ve loosely mentioned agents, perhaps I’ve used device and agent interchangeably. An agent has some sensors, a communication mechanism and actuators through which it might sense and manipulate its environment and communicate the results. (Only the first two are crucial, agents without means of modifying their external environment I will call passive.) In the case of the electric grid, an agent is some device M (i.e. electric vehicle, substation, storage unit, etc.) that has a program l running on it, so that it is $agent(l, M)$. Be aware that my examples are in the context of a power system, although I frequently personify inanimate assets for the sake of explanation.
When multiple agents interact to solve problems that are beyond the capabilities of a single agent, it is referred to as a multi-agent system (MAS). Multi-agent systems are <u>everywhere</u>. They best describe human social interactions and markets.
I have said a lot at this point. It is my hope that this gives the impression of an ad hoc market driven agent-network; which owes its flexibility to impromptu and direct messaging between devices. A Wild West of self-serving grid-agents run amok sounds risky. (These agents are selfish but not inherently competitive and biased towards self-organization). Therefore, MAS based power systems should be self-governing, comprehensible and transparent – agents should be capable of justifying themselves when interrogated. Note that scrutability/transparency does not guarantee an agent’s predictability; what really matters is that hindsight is 20/20, foresight might well be nonexistent. (A leap of faith we frequently find ourselves in due to the ubiquity of black box artificial intelligence.) Of course that is not to say we should not or could not predict an agent’s behavior, all I am saying is that we cannot underestimate a power system’s complexity and overestimate an agent’s rationality. I later propose means of prediction and grid diagnostics. Ideally, a grid-level supervisory layer would use existing infrastructure and maintain order and grid-wide monitoring <u>without</u> interfering in the daily operations of private agents; a sort of robotic laissez-faire. Agents participate in *optimistic transactions*: when aware of an existing or impending market failure, an agent will select a compensating/corrective action. Whether a designated group of

agents (read: supervisory layer) yields a *fairer* network than all agents having the capacity for altruism requires additional research.

** Insert figure of power system casual operation, and then the monitoring for utilities - define how private agents communicate using standardized protocols that a supervisory agent can then interpret **

Let us develop the MAS further. A critical aspect is giving 'dumb' assets on the grid agency. By this I mean autonomy, a catch-all term for a limited understanding of causality and the ability for agents to coordinate themselves. Some reasons for this are to: (1) enable more sophisticated demand side response (DSR)[14], (2) enhance grid awareness and observability, (3) improve robustness for the safe integration of intermittent or misbehaved assets *plug-and-play*, and (4) enable more inclusive and personalized markets. These aims are neither new or surprising yet I am often asked however as to what I mean by causality in the context of a power system so allow me to clarify. Suppose an autonomous DER[15] discovers its next-day bids are all rejected by the market maker. Reducing its prices appears to have little to no effect on its chances of selection. (The agent runs several varied experiments and statistically arrives at this conclusion). As to why the rejection, assume the *human* market maket does not believe that the DER in question can deliver the allotted amount of power the next day due to conflicting weather reports; further, for lack of rapport, the operator doubts the DER could pay the financial penalties for erratic behavior. How to 'relay' such a line of reasoning to a machine? Perhaps it need not be relayed at all, the agent may discover this on its own: one method is via 'consultation'. The DER confers with neighbors of roughly the same type. After comparing itself to its successful and unsuccessful peers, it may independently hypothesize a link between bid failure and unreliability – a notion of unreliability did not originally exist in the agent. The DER, in response to this abstracted notion of unreliability, may decide to partner with underutilized natural gas generators to form a stable hybrid group[16]– responsibility is now distributed. In this scenario, causality is the understanding that to every effect (rejected bids) is generally some cause (unreliability). So long as the operator is somewhat consistent in their decisions, an agent can hypothesize the underlying cause and effects via the comparisons proposed or another suitable mechanism. Simple dichotomies like acceptance and rejection, or swiping left or right, or liking or disliking are powerful learning tools for agents conscious of it. Bayesian inference could also be applied here. As more system interactions emerge, MAS may enable grid operators to confidently catalog grid events knowing that conflict-resolution is being done in the background.

To tie things together better, I introduce the *private problem* and the *collective problem*. The former describes the tasks whose completion are the reason for the agent's 'being'; i.e. a private PV system is bought and created so that it may provide a building with a means of power outside the grid. I pose a question to initiate discussion on the 'collective problem': can high levels of grid observability be achieved through the analysis of interactions between potentially disinterested agents? (Disinterested here means an agent's primary and 'exclusive' interest is the satisfaction of its owner). Like Bitcoin, one could incentivize efforts towards the collective problem via financial compensation, but this can lead to perverse effects. i.e. a storage unit programmed to maximize profits for its owner may discover contributing to the collective problem (or not contributing at all) to be more lucrative than fulfilling its primary purpose of the provision of stored power (or helping maintain the environment in which it otherwise profits – free-rider problem). Either way, a MAS platform should be agnostic to the inherent cost functions of agents and should avoid hi-jacking the computing resources of independent agents under the assumption that agents are selfish. (These requirements would preempt the heavily distributed ownership and siloing I anticipate will steadily characterize the grid). In more familiar terms, the goal of a capitalist society is the transmutation of the egoist nature of man into favorable outcomes for the society. If you prefer, the initial question can be phrased differently: can cultural, social and economic claims about a society be made solely through observation of the interactions of its inhabitants without resorting to personal interviews and interrogations? Or better yet, can you divine the events transpiring within an ant mound without dismantling it? Formally, a system is observable if for any permutation of its possible states and control vectors (i.e. agent and the actions felt by the grid), the current state can be determined in finite time using only the system's outputs.

Back to the question at hand. Causal evidence suggests yes. (Network inference has been applied to epidemiology: using the sequence of node infections in a

[14] DSR is the manipulation of user energy demand through various signals and controls. A 'primitive' form would be randomly disconnecting loads during times of peak energy demand.

[15] An agent without a human controller.

[16] Pairings between renewable power producers and underutilized fast-ramping natural gas plants are explored in this paper [6] and are called reliability contracts.

graph, (*traces*), and a probabilistic model for the epidemic [3]). I suspect that the data characterizing agent interactions will possess more useful meta data or be equivalent to that generated by SCADA. If an engineer can derive insights from there, why couldn't a machine (or machines in our case)?
Let's further develop the supervisory layer touched on previously.

2.1. Supervision and Distributed State Estimation (SE)

There is no current model for state evolution – the grid being affected by many factors at the same time. As I know, a catalog/knowledge base detailing grid events does not exist. (Perhaps power system engineers and operators have an intuition for grid evolution that comes with years of on the job experience.) This limits the methods of estimation that can be used. Plugging in local measurements into local estimators and passing it on to a central coordinators for further estimation is an example of one such method – hierarchical. With an increase in the number of measurements and sampling rate, it is to be expected that communication bottlenecks and reliability become issues for hierarchical SE. What we need is a fully distributed SE method, the requirement being that the communication graph is connected. (The communication graph does not need to overlap with the physical T&D network – information can be transmitted in a myriad of ways).
Traditional graph theory has an undirected graph given by $G = (V, E)$ where $V = [1, \ldots, N]$ is a set of nodes (agents) and E is a set of edges of which there are $|E| = M$. The unordered pair (n,l) is in E if agents $n \in V$ and $l \in V$ share an edge. The neighbors of agent-n are $\Omega_n = \{l \in V | (n,l) \in E\}$. The degree of agent-n is $|\Omega_n|$, the number of edges connected to it. The actual communication graph's structure is an adjacency $N \times N$ matrix $C = [C_{nl}]$ in which $C_{nl} = 1$ if $(n,l) \in E$ and $C_{nl} = 0$ otherwise. We may construct the degree matrix D whose diagonals are the degrees of agents-1 to agent-N, and from there give the Laplacian matrix:

$$L = D - C \tag{3}$$

Which tells the following:

$$L_{nl} := \begin{cases} |\Omega_n|, & \text{if } n = l \\ -1, & \text{if } n \neq l \text{ and } (n,l) \in E \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

From this is a notion of the network's synchronizability. The *Fiedler vector* can be used to partition the network into clusters. A similar construction can be made for the physical T&D network – it can be complemented by existing physical asset data files (GIS) and models. Clever partitioning of a system could give insights into how to best route information/power. Something I considered in the past was whether grids could be sustained using only renewables without any storage; intermittency was to be mitigated by increasing the network's connectivity and size, *it is always sunny somewhere*, and exploiting seasonal and regional disparities in the availability of electricity. I most thought about the stable configuration of the network. An ideal configuration would take into account the size and density of interconnects and the capacity and dispersion of renewables among other variables. Although my results were inconclusive, I have come to appreciate network topology's impact on operational efficiency.

Many redundant observations of power flows within a network can yield the voltage phasors of the system's buses at some instant. The step in between applies inference. The underlying state of some network X may be described by the voltage magnitude and angle at each bus except the reference bus. ($\theta_0 = 0$ at the reference).

$$X = [V_0, V_1, \ldots, V_N, \theta_1, \ldots, \theta_N] \tag{5}$$

A state estimator (SE), which really is an agent, observes $H_k\theta_k(i) + \eta_k(i)$ at discrete time indexed by i. H_k is an observation matrix and $\theta_k(i)$ is agent-k's state vector, which includes its local voltage-angle measurements. The noise to each measurement follows a zero-meaned Gaussian distribution such that $\eta_k \sim N(0, \sigma_k^2)$. The deviations in error are known for each agent. To construct a fused state estimate $\bar{x_n}(i)$, agent-k fuses its own estimate $x_n(i)$ with its neighbors Ω_n to obtain $\bar{x_n}(i)$. This fusion may be the average of its neighbors much like the centralized least squares approach. [13] is a distributed algorithm that uses data exchange between smaller control areas (with the loose requirement that $C_i \subset N$ where $\forall (n,l) \in C_i$ there exists some sequence of agents that creates a physical path between n and l: the definition of connected) to construct a global estimate $\hat{X}$. See [13] for more.

It is desirable to come up with representations for the grid because agents interact through it. (For most cases, a power system agent cannot affect an agent directly). The *Tree Model* I propose captures some of the assumptions I make about the grid's responsiveness and behavior. The model involves as of now naive algorithms. Further refinement is needed. To start with, the model is referred to as a tree because there is a *trunk* and many branches. At the end of each branch is some collection of agents. Keep this mental image in your head. I claim that the grid's behavior at time t_1 is determined by all agent actions at t_1 and its past behavior. The current state depends on past states $X(x_n|x_{n-1}, x_{n-2}, \ldots, x_1)$[17]. This captures the *lollapalooza* effects that I think are present on the grid. I make some assumptions: (1) at any one time, I claim the grid resists anything that seeks to modify its trajectory. I can make a data structure, the *trunk*, to describe this phenomenon, (2) with that said, coordinated actions by agents can overcome the grid's resistance and alter its trajectory. Spatial proximity affects the effectiveness of coordination (a suitable assumption considering power dissipation effects), and (3) the informational content of power flows is communicated to other parts of the grid via the grid's trunk. Now is the time to make the static model dynamic. Each agent can oscillate up/down, right/left, or stay still. Now, the tree is in a way 'rustling' – disturbances at the branches cause torsion in the trunk. (The trunk is flexible). When the agents oscillate forcefully, or in a manner that matches the trunk's natural frequency, the trunk movements become chaotic and it eventually breaks.

I (rather vaguely) say the grid at time t is in a state $x \in \{X\}$, and has n independent agents who each possess a set of actions $A = \{a_1, a_2, \ldots, a_b\}$. (Agents can execute only one action at any time). There exists a transition function $T : X \times A_1 \times \cdots \times A_n \to PD(X)$ whose transitions are controlled by the current state and an action from each general agent A_i. (I use similar notation to a Markov game[18] described in [1] albeit I earlier disagreed with the Markov property. If you grow annoyed by my casual misuse of Markovian notation when I have stated my disagreement with its defining characteristic, you may think of a recurrent neural network (RNN) formulated here [11]). This can also be written compactly as $T(x, A', x')$ where $x \in \{X\}$ is the current state, x' is the next state and A' is a collection of single actions from each agent. Assumption (2) which says that there is an A' such that $T(x, A', x') \approx 1$. In an electric grid, agents that are close to each other may form a bloc, so realistically for a given topology, patterns within A' may suggest the locations of coalitions or individual agents. For example, say a grid is made of five agents with actions a and b and the following A''s are observed to frequently lead to state changes $[a,a,b,b,a], [b,b,b,a,b], [a,a,a,b,b]$ such that for each $T(x, A', x') \approx 1$. Additionally, say the grid's state may be driven to a or b. Comparing the A''s and following assumption (2), we could hypothesize that agents-1, 2, and 5 are on average closer to each other than they are to agents 3 and 4 because when that particular set of agents are coordinated, there is a higher probability the grid's state changes. The challenge lies in discovering the transition function which is a proxy for the system's underlying dynamics. A transition function abstracts away some of the physical minutiae and hopefully provides interesting insights.

Say we have actions up, down, agents A, B, C, D and the 3x3 topology looks like the following:

Figure 1. Grid 1

A	B	
C		
		D

up	up	
up		
		down

In Figure 2.1, it is clear that the bulk behavior of the system is driven up. If we change the topology by adding another agent E at position (2, 2) as in Figure 2.1, it becomes less clear what the system is driven towards. One could claim that the blue region yields nothing, to give the effective topology on the right.

Figure 2. Grid 2

up	down	
down	up	
		down

		down

[17] Although the grid-state is analog, it can be made discrete. Also, note I violate the Markov property which states that a system's future and past states are independent. I believe, perhaps incorrectly, that past states of the grid may persist and constrain the grid's future behavior – and so there is memory. Note I have come across conflicting claims that everything is Markovian and that nothing is truly Markovian.

[18] A reader familiar with Markov games would be dismayed that I make no mention of reward. Self-admittedly, I do not know how to design an appropriate reward mechanism that unifies opportunistic agents, for which immediate feedback by way of a reward is sensible, versus agents with some notion of grand strategy. Also, how to assign reward to human-like behaviors (which I expect to see in the MAS) that we ourseves find morally amibiguous? I sense that there is an easy fix, but I do not understand *game-design* enough to think of it. For now, concern yourselves only with the state-transition aspect of Markov games, in later sections I address how agents might act.

2.2. Task Allocation and Objective Signals

To achieve some objective O, a set of agents N operate in strictly one of two modes. An agent can work independently on personal objectives (i.e. voltage regulation, profit maximization) or within a coalition $C_i \subset N$. The former mode of operation is to be proactive, while the latter reactive. Varied levels of control are desirable for a utility. As an example for the latter, the utility dispatches the load-shed signal through the network with some total reward. (The load-shed signal for simplicity is formatted using the same primitives later developed in the **communication** section). Prominent agents break down objective O into a set of tasks $T = \{t_1, t_2, \dots, t_m\}$ with an optional priority order $t_1 \preceq t_2 \preceq \dots \preceq t_m$. (Resource allocation is a means of prioritizing tasks, such as a restriction in communication to permit collaboration between the relevant agents). It then advertises tasks to nearby load and generator agents. Agents who accept the request are made members of the coalition C_i. Rewards are disbursed retroactively to make the process quicker. The reward mechanism can be improved.

** Elaborate **

An additional challenge lies in breaking down some objective signal (assumed to be a string) into a set of feasible tasks. (An objective could be described in a way that is already decomposable). Pattern based schema and natural language processing may be useful here.

** Enter any insights into how to break down a string into its components - I have a hunch that humans work via comparisons and proportions. **

Now, how to make an agent intelligent?

2.3. Intelligence and Reasoning

There are many ways to imbue an agent with 'intelligence'. The entire field of artificial intelligence is dedicated to it!

2.3.1. Weak Approach, Using Fixed Deduction Rules ρ. Let L be the set of all possible observed interactions in an environment, i.e. *agent-c observes that agent-a bought x kWh at price y USD from agent-b*. A is the set of all actions available to the agent. The database of all observable interaction groupings would then be the power-set of L, $D = \mathscr{P}(L)$[19]. That's to say if $L = \{x, y, z\}$, with x, y, z being random interactions, then,

$$D = \{\{\}, \{x\}, \{y\}, \{z\}, \{x,y\}, \{x,z\}, \{y,z\}, \{x,y,z\}\} \tag{6}$$

The catalog of observations (read: internal state) of an agent-m is then $\Delta_m \in D$. To determine which action to choose according to some encoded deduction rules ρ (rules the programmer implements in the agent) I present the following pseudocode wherein the chosen action is defined by agent's internal state.

Algorithm 1: Determining actions via hard-coded deduction rules

input : Global database D of observable interaction groupings
output: Agent action

```
Function action (Δ : D) : A
    for each a ∈ A do
        if Δ ⊢_ρ Do(a) then
            return a;
        end
        return none;
    end
```

For every possible action, if the action in question is derivable from the agent's internal state via the hard-coded deduction rules, then return that action. Quite simply, if Δ were a vector and $a \in A$ a scalar, action is a mapping $action : \Delta \rightarrow a$. There are some issues with this approach. Some of which are an over-reliance on calculative rationality [12], an inability for an agent to learn if ρ is static, and the difficulty in constructing ρ. Also, what happens when None is returned? These are as of now unanswered questions.

2.3.2. Other. Some other learning algorithms might include neural networks and reinforcement learning. For the former, minimizing *catastrophic forgetting* (the tendency for an agent to forget how to perform task A after learning task B) in face of unpredictabile task sequences is paramount. [5] is excellent in that it describes giving weights in a neural network an elasticity that makes them resistant to change to replicate studies of performance retention in mice[20]. Research into con-

[19] It might be more accurate to say that $D = L^*$, the set of finite-length strings generated by concatenating elements of L; the same element may be used multiple times.

[20] "When a mouse acquires a new skill, a proportion of excitatory synapses are strengthened; this manifests as an increase in the volume of individual dendritic spines of neurons [Yang et al.,

tinual learning is exciting, and I will share my comments in a later paper. In an earlier example regarding a DER and natural gas power producer partnership (inspired by [6]), I proposed agent reasoning achieved via consultation with peers. It may be possible that learning/reasoning may also be facilitated by narratives [21]. A narrative would be a data structure of temporally ordered events centered around the agent in question whose creation allows the agent to achieve coherence (read: belief revision) in its strategy and to transmit it to fellow agents. From narratives, peers can determine the degree of relatedness of desires/intentions between themselves and another agent which may inform its decision making.

2.4. Agent Negotiation/Communication

2.4.1. Negotiation. *Negotiation: 'A process by which a joint decision is made by two or more parties. The parties first verbalize their contradictory demands and then move towards agreement by a process of concession making or search for new alternatives.'*

During a negotiation, an agent would:

1. Evaluate incoming arguments made by its opponent and update its mental state accordingly. For example, if it is communicated to agent-a by agent-b that the voltage/current sensed by agent-b has risen, then agent-a may shed some load downstream to avoid voltage disruptions.[22]

2. Generate candidate outgoing arguments.

3. Select an argument from the set of candidate arguments. Agents may assign to each candidate argument some *convincing factor* a value between 0 and 1 that denotes the agent's confidence that its opponent will accept it. (Agents are designed to be persuasive). The agent would also vet the candidate arguments for conflicts with its own desires.

Now, as to signaling on the grid. How should an agent inform another of its intentions, beliefs and desires?[23] Prices are generally the go-to means for capturing the pertinent information. (Locational marginal prices in theory reflect the network's constraints, available generative resources and all other relevant demands). I am skeptical however. Prices can be obfuscating – they fail in giving the reasoning behind actions, signaling only how much "someone desires something relative to someone else"; without justification or for fears of assuming intentions incorrectly, why should anyone change their desires, intentions and style of negotiation? I hope it is somewhat apparent my issue with taking prices prima facie. Because I am realistic however, I understand the challenges of assigning value to things in a way that is communicable, useful and accurate. This is why I proposed the external analysis of agent interactions as a method for extracting more meta-data. (As we shall see, agent communications are designed in a way that more information is transmitted than just message content.) This is not altogether dissimilar to a neural network's ability to detect deeper structure in data otherwise invisible to us.

2.4.2. Communication. Is it possible to optimally coordinate the actions of agents <u>without</u> an overlying communications network? There are examples where the electrical coupling between agents is exploited for active control purposes, i.e. as we have seen, load sharing and synchronization can be achieved naturally in electrically coupled synchronous generator systems. These examples appear to only be applicable where the physical dynamics are well understood[24]. This section is for the converse cases or those where agents cannot readily be made to emulate well-understood dynamics. Communication topology may be different to the physical topology for power interconnections. A challenge to overcome is obtaining the bandwidth and speed necessary to make inter-agent communications meaningful. Not to mention that information fidelity and its impact on a power system's controllability is not well understood. We can make qualitative assertions like frequent packet dropping negatively affects a system's controllability, but cannot say the degree to which it harms performance. Power line communication is a 'through-the-grid' communication technlogy which uses power lines as a medium for information transmission. Interestingly, [2] describes how grid topology can be inferred through PLC; which complements grid diagnostics and the supervisory layer discussed earlier. (Note that the grid was not designed to support high-frequency signal propagation and that PLC is susceptible to considerable signal attenuation and reflections

2009]. Critically, these enlarged dendritic spines persist despite the subsequent learning of other tasks, accounting for retention of performance several months later [Yang et al., 2009]. When these spines are selectively "erased", the corresponding skill is forgotten [Hayashi-Takagi et al., 2015, Cichon and Gan, 2015]" [5].

[21] A notion somewhat inspired by [7].

[22] This of course assumes agent-b can be trusted.

[23] A reader familiar with MAS might recognize this as the BID methodology.

[24] I have heard people flippantly say that we know everything there is to know about electromechanical generators. Always the skeptic, I am dubious of such claims but must admit the field is well studied.

from line discontinuities).

** Elaborate **

Everyday agent communications may be conducted using the formal logic described by [8]:

$$\neg[t_1, Do(i,\alpha)] \rightarrow [t_2, Do(j,\beta)] \tag{7}$$

Here, Agent-j threatens to perform action β at time t_2 if agent-i does not perform action α at time t_1. Note how the ordering and qualifiers enhance the informational content. The intention of such logical descriptions is so that primitives may be extended further and further. So long as all agents have access to the same primitive build-set, they should theoretically be able to interpret fairly advanced locutions. Argumentation, like negotiation, would emerge from exchanges between persuasive agents whose aims are to bring about changes in intentions of their opponent. It is through the aforementioned locutions that agents impart information to one another – which may complement or substitute an agent's inference system, which may be weak or bounded. The following is a dialogue between a buyer (B) and seller (S) taken directly from [9]. For context, the buyer would like to rent a car.

B: I would like to rent a car for four days please.
S: I offer you one for $400.
B: I reject! How about $200?
S: I reject! Why do you need a car?
B: Because I want to get to Sydney.
S: You can also go to Sydney by flying there! I can book you a ticket with Qanta airlines for $200.
B: Great, I didn't know flights were so cheap! I accept!

The dialogue demonstrates how inquiries, which may be phrased using the same logical framework above, can lead to agreements that satisfy both a buyer and seller. Agents may help each other using referrals (and be compensated for it) and facilitate the *societal benefit*.

Repeated encounters between agents can of course lead to learned (albeit abstracted) notions of credibility and reputation. When combined with the narrative structure discussed previously, agents may maintain durable profiles of other agents that persist if memory is sufficient. I am currently looking into mechanisms where agents simulate the content of a locution within sand-boxed environments. (These environments are built using the agent's mental state.) Another interesting communication protocol is contractnet as described by [4]. The paper suggests agents be given roles of manager and contractor *a priori*. (The roles are dynamic, an agent can only be one per negotiation). Here is the recommended format of [4] for communication:

Figure 3. ContractNet Proposal

To: * indicates a broadcast message.
From: 25
Type: TASK ANNOUNCEMENT
Contract: 22-3-1
Task Abstraction:
TASK TYPE SIGNAL
POSITION LAT 47*N* LONG 17*E*
Eligibility Specification:
MUST-HAVE SENSOR
MUST-HAVE POSITION AREA *A*
Bid Specification:
POSITION LAT LONG
EVERY SENSOR NAME TYPE
Expiration Time:
28 1730Z FEB 1979

For context, this is the message a processor node would broadcast to relevant sensors to help it in constructing a map.

My only disagreement with this is that it necessitates hand-engineering instructions/commands and is degraded when there is an interaction between tasks. (Recall the objective signal O broken down into tasks $T = t_1, t_2, \ldots, t_m$, interactions could be $t_1 \iff t_n$, which reads "t_1 is feasible if and only if t_n is"). That is not to say I disagree with applying domain-knowledge. Agents will not be as smart. As for the communication protocol used in the later **agent demonstration** section, I decided to keep it simple using a publish-subscribe pattern and the proposals are communicated using dictionaries whose key is the issue at hand, and value is the issue's scalar value. (I am currently implementing logical programming of the sort characterizing the locutions).

2.4.3. Security. The issues of security covers a field in which I have no special competence. Therefore, the following is a cursory survey of security vulnerabilities. While distributed approaches like this MAS in particular remove a single point of failure, the increased complexity invites more points of attack. The device-agnosticism of the proposed platform is also liable to be a security issue. This vulnerability is particularly concerning because current grid communication protocols with similar agnostic-design in mind (Modbus) do not support authentication and confidentiality. I think selective-endorsement as a means to generate *inter-agent trust* can be appropriately applied here. When an agent joins a network it is to be endorsed by other well-trusted agents - different parts of the network have different trust requirements in terms of the number of agents needed to endorse another agent's entry. Intruders discovered to be malicious are ejected and banned from the network – work may go into reversing the commitments made by the malicious agent and defining the endorsement protocols. (i.e data integrity/verification network protocol like check-sums can verify agent identities and actions).

2.4.4. Generalized Bilateral Negotiation Model. Consider a two-agent system, composed of agent-a and agent-b. (Agent-i $\in \{a,b\}$). Each general agent-i has $J = \{1,\ldots,n\}$ the set of issues it would like to negotiate under a contract, and the acceptable value x_j for issue $j \in J$ that it might accept is in the range $x_j \in [min^i_j, max^i_j]$. Agent-i scores an issue j according to the scoring function $V^i_j : [min^i_j, max^i_j] \rightarrow [0,1]$ - scores are kept within the $[0,1]$ interval for simplicity. Furthermore, agents assign importance to each issue j with normalized weights such that $\sum_J w^i_j = 1$. The scoring for an entire contract composed of many values $(x_1,\ldots,x_n)$ is then:

$$\sum_J w^i_j V^j_i(x_j) \tag{8}$$

For instance, a seller agent, (agent-a), may have negotiation issues price, volume, duration. Agent-a could then have a price (USD) reservation within $[10,20]$, volume (kW) reservation within $[1,5]$, and preferred contract duration (hours) within $[3,7]$ with greater emphasis on price, then volume, then duration described by weightings $(0.5,0.3,0.2)$. To each issue, agent-a has linear scoring functions:

$$V^a_{price}(x_{price}) = \frac{x_{price} - min^a_{price}}{max^a_{price} - min^a_{price}} \tag{9}$$

$$V^a_{volume}(x_{volume}) = 1 - \frac{x_{volume} - min^a_{volume}}{max^a_{volume} - min^a_{volume}} \tag{10}$$

$$V^a_{duration}(x_{duration}) = 1 - \frac{x_{duration} - min^a_{duration}}{max^a_{duration} - min^a_{duration}} \tag{11}$$

When agent-a receives multiple contract offers, it rationally selects the one with the highest overall value. Suppose three buyers, agent-b1, agent-b2 and agent-b3 submit the following offers respectively to agent-a (seller): $[11,4,4], [16,5,7], [10,2,3]$ – recall the ordering is [price, volume, duration] and agent-a's reservations are $[[10,20],[1,5],[3,7]]$ and weightings are $[0.5,0.3,0.2]$. Given the above parameters, we can compute the value of each offer:

Figure 4. Contract Example

Agent	*Offer*	*Calculations*	*Offer Value as seen by agent-a*
B1	[11, 4, 4]	*price*: $\frac{11-10}{20-10} = 0.1$ *volume*: $1 - \frac{4-1}{5-1} = 0.25$ *duration*: $1 - \frac{4-3}{7-3} = 0.75$	$0.5 * 0.1 + 0.3 * 0.25 + 0.2 * 0.75 = 0.275$
B2	[16, 5, 7]	*price*: $\frac{16-10}{20-10} = 0.6$ *volume*: $1 - \frac{5-1}{5-1} = 0$ *duration*: $1 - \frac{7-3}{7-3} = 0$	$0.5 * 0.6 + 0.3 * 0 + 0.2 * 0 = 0.3$
B3	[10, 2, 3]	*price*: $\frac{10-10}{20-10} = 0$ *volume*: $1 - \frac{2-1}{5-1} = 0.75$ *duration*: $1 - \frac{3-3}{7-3} = 1$	$0.5 * 0 + 0.3 * 0.75 + 0.2 * 1 = .425$

By the computations above, agent-a should accept agent-b3's bid. Why b3 over the other agents? The answer lies in the value functions. The value functions by their linear definitions reward high prices and penalize high request volumes and durations. (A seller would aim to sell as little quantity as possible at the highest price possible to any single buyer. The little quantity is counter-intuitive but makes sense when we consider that a seller has limited volume and would prefer to disperse their product as widely as possible to avoid dependency on a few buyers. One could also assume our agent-a cannot provide an incremental amount of goods without enduring rising marginal costs, i.e. agent-a is small and cannot exploit economies of scale). As seen from the offers, b3 offers the lowest price, but also accepts the lowest volume and time duration of the three; this allows it to just outbid the other agents. B2 is close, it is willing to give a high price but suffers from asking the maximum volume and duration – if b2 were to downgrade its volume request from say 5 to 3 kW for an unchanged 7 hours, then its offer value would be 0.45 and beat out b3's offer. For similar reasons, b1 has the lowest valued offer: it asks for too little in price, and relatively too much in volume and duration.

There are some failings with the model above:

1. *Lack of consideration for resources.* Particularly with energy, selling energy at some instant deprives me of it at a later time. This is especially true for prosumers (those entities that both sell and indulge in their own product, agents/prosumers = buyers $\cup$ sellers). There

should be a measure for the anticipated use of energy or a sense of urgency. If agents were greedy and opted for maximizing utility at each turn, performance would suffer for this reason.

2. *Transparency of information*. The transparency of the scoring functions and reasoning models are inappropriate for competitive negotiations.

3. *Assumptions of behaved value restrictions*. Restrictions on value here are artificial. In many cases, negotiation involves determining whether regions of intersectionality even exist.

3. Agent Demonstration

To make discussions up to this point concrete, I provide a demonstration of a two-agent MAS that has been coded as a Python simulation. (In writing the program, I used an object-oriented approach but was conscious that agents and objects differ in that agents exist within their own encapsulated containers while objects exist within the same 'space'. The simulation has some of the behaviors to follow already implemented). The participants include a charging system (CS) and an electric vehicle (EV) – both desire the EV charged at acceptable issue values. The EV submits a charging request with an initial offer to the CS, to which the station responds with its own offer as determined by the station's own state: current (and forecasted) idle ratio of charging devices, load of the local power grid, wiggle-room for any financial penalties, quantity and quality of dispatchable renewables, etc. The EV also has some initial expectation of the price. Note that this current model is price centric. With some adjustment, it may be generalized to include other issues. Note for this section, I am grateful to the authors of [14] for the following price formulation sections.
The 24-hour day is divided into 10-minute timeslots. Whenever charging price negotiation is formed within a timeslot, the EV and CS may agree on whether it is worth renegotiating the contract for forthcoming timeslots, or to leave the contract as is.
I will be applying some of the concepts introduced in the *bilateral negotiation model* described in an earlier section.

3.1. Initial Price Formulation EV

The following factors are considered by the EV_i in its expectation of the offer returned by the CS_j.

1. Traveling requirement of EV_i, R_i, in the next period n – characterized by,

$$R_i = X_{i,n}\frac{\lambda_1}{n}(\lambda_2 M_{i,n} + \frac{\lambda_3}{SOC_i}) \tag{12}$$

$$X_{i,n} = max(0, \text{ plan to travel in n?}) \tag{13}$$

$M_{i,n}$, SOC_i are the mileage requirement for the next period and the current State of Charge of EV_i. The lambda coefficients denote the influences of the urgency of travel, mileage requirement and SOC.

2. The grid's current load $L(t)$. As $L(t)$ increases, as should the EV's psychological expectation of higher prices.

3. The distance $d_{i,j}$ between EV_i and CS_j. Charging convenience sensibly decreases as $d_{i,j}$ increases and so represents an inverse relationship between expected price and convenience of distance.

The initial price an EV_i with a baseline price of p_0 set by the owner, might expect would then be defined by:

$$p_{i,0} = p_0(R_i + \gamma_1 L(t) + \frac{\gamma_2}{d_{i,j}}) \tag{14}$$

The gammas indicate the influences of current load of the power grid and the distance between the EV and charging station respectively.

3.2. Initial Price Formulation CS

CS_j will formulate a price that maximizes its own interests. To that end it considers:

1. The ratio of idling to charging devices in use, IR_j:

$$IR_j = \frac{CP_{inuse,j}}{CP_{total,j}} \tag{15}$$

Where CP corresponds to the charging devices in use.

2. The current load $L_j(t)$ versus the total available power it could call upon. The latter includes things like renewables and coordinated power arrangements with third parties.

$$L_{j,0}(t) = L(t) + \sum_{i \in N} P_i \tag{16}$$

$$P_i = \frac{(A_i - SOC_i)Q_i}{t_{i,0}} \tag{17}$$

$L(t)$ is the load with no EV's attached, N is the total number of EVs and P_i is the power delivered to EV_i. A_i is the charging threshold of SOC (can be the EV's target level of charge), Q_i is the EV's battery capacity and $t_{i,0}$ is the forecasted charging time.

The charging station operates under the constraints that current load does not exceed its maximum rated load $L_{j,max}$. Also, EV_i's charging power P_i must exist between the charging station's min and max charging power and that of the EV; $P_{MIN} = max(P_{i,min}, P_{j,min})$ and $P_{MAX} = min(P_{i,max}, P_{j,max})$ so that $P_i \in [P_{IN}, P_{MAX}]$. The charging time should not exceed the EV's requested charging time t_i ($t_{i,0} \leq t_i$), although that can be left to negotiations.
The initial charging price is then:

$$p_{j,0} = p_0(1+\delta)\left(\alpha_1 IR_j + \alpha_2 \frac{L_j(t)^2}{EL_j}\right) \tag{18}$$

Where EL_j is the average daily load of CS_j and δ is the expectation of profit. The alphas indicate the influences of the idling charging device ratio and current load. Charging price increases with the square of 'discounted' current load to achieve peak shaving. Do pay attention to the variables like gamma and alpha. These values can be adjusted as the agent learns through its negotiations.

3.3. Termination Rules

A transaction is reached through a multi-round price negotiation thread, either until an agreement is made, or a party abandons the negotiation. To encode the latter, each agent has some K[25] which is the round number they are willing to reach until opportunity cost causes them to exit. Each agent has three choices – for EV_i:

1. $p_{i,k} \geq p_{j,k}$, agreement is reached, transaction goes through.
2. $p_{i,k} < p_{j,k} \wedge k < K_i$ then EV_i does not accept the price but is willing to negotiate into the next round.
3. $p_{i,k} < p_{j,k} \wedge k \geq K_i$ then EV_i abandons the negotiations with CS_j.

For CS_j:

1. $p_{i,k} \geq p_{j,k}$, agreement is reached, transaction goes through.
2. $p_{i,k} < p_{j,k} \wedge k < K_j \wedge L_j(t) \leq L_{j,max}$ then CS_j refuses the price but is willing to continue negotiations into the next round.
3. $p_{i,k} < p_{j,k} \wedge (k = K_j \vee L_j(t) > L_{j,max})$ then CS_j may unilaterally terminate the price negotiation.

[25]The values of K will be informed by some proxy measurement for opportunity cost. Agents can't negotiate forever and so K could also be a measure of time.

3.3.1. Price Adjustment for CS_j.

CS_j could compute the average bidding price of all its current bidding EV users for the last k-round, $\overline{p_k}$. CS_j would also compute the average bidding users for the same time, $\overline{N_n}$ – wherever $N_n \leq \overline{N_n}$ then CS_j may choose to reduce the price, otherwise, the charging station may increase the price. Next round's price is then determined by the ratio of the current to average number of bidders (supply and demand), and the average bidding price.

$$p_{j,k+1} = p_{j,k}\left(1 - \omega(\ldots)\frac{\overline{N_n}}{N_n}\frac{p_j - \overline{p_k}}{p_j}\right) \tag{19}$$

ω is a decision-making function. When $\omega(\ldots) = 0$, then CS_j wants no price reduction.

3.3.2. Price Adjustment for EV_i.

From CS_j, EV_i will receive the average bidding price $\overline{p_k}$. Whenever EV_i's price is rejected, the EV will raise its price proportional to its urgency. The next price will have the form:

$$p_{i,k+1} = p_{i,k}\,\omega(\ldots)\left(1 + min(0.05, \frac{\overline{p_k} - p_{i,k}}{p_{i,k}})\right) \tag{20}$$

$\omega(\ldots)$ is a tactic function described in more detail shortly.

3.3.3. Tactics for Changing Variable k+1 Price Adjustment.

Tactics are functions that affect the change of an issue's valuation. Generally, counter proposals are affected by the weighted combination of different tactics over a set of issues. Tactics are applied to the price change functions multiplicatively as per $\omega(\ldots) = \sum_{n=1} \mu_n \omega_n$, a function composed of tactic functions $\omega_1, \omega_2, \ldots, \omega_n$ that depend on as of yet undetermined quantities; μ_n are variables. In the simulation we present later, I choose to simplify the application of tactics, and so the price change function is affected only by a single random tactic at a time.

1. *Time and round dependent*. When there is a deadline, tactics ought to model the fact that an agent will concede more rapidly as the deadline approaches - the curve of concession, a time dependent function, differs between agents. For the sake of simplicity $t_{max} \equiv K$ and $t \in N^+ \equiv k$. I select an exponential function $\omega_{time} : t \rightarrow [0,1]$ whose degree of convexity is parametrized by $\beta \in R^+$ to be the concession function. The functions behavior is initially constant, then as the time deadline nears, it increases. The function is of the form:

$$\omega_{time}(t) = exp\left(c\left(1 - \frac{min(t, t_{max})}{t_{max}}\right)^{\beta}\right), c \in (0,1) \tag{21}$$

c is determined by the initial expectation of price of the agents. A higher c generally means quicker capitulation. When I let $t_{max} = 1, c = ln(0.5)$ and vary β I get the following curves:

- *Take it or leave it*. When $\beta < 1$, the function only ups the value to its reservation when time is nearly exhausted.
- *Concessive*. When $\beta > 1$, agent quickly meets its reservation value.

2. *Imitative*. Whenever an agent is under no pressure to reach an agreement, its counter proposals may be informed **only** by its negotiation opponent. That is, the agent may imitate certain aspects of its opponent's behavior. For example, a charger may decline its selling price by 5 USD if a potential EV increases their willing price by 5 USD; this is a case of *absolute Tit-For-Tat*. Imitative behavior is applicable only when there is a negotiation history with the opponent.

   - *Absolute Tit-For-Tat*. Agent-i[26] mimics opponent agent-j's behavior some δ steps ago. A random quantity is included to escape any local minima and advance negotiations.

   $$\omega_{A-tit-tat}(p^i_{k-1}) = 1 - \frac{p^j_k - p^j_{k-\delta} + (-1)^s R(M)}{p^i_{k-1}} \tag{22}$$

   $R(M)$ returns a random number in the interval $[0, M]$, M being the maximum amount an agent is willing to deviate from pure imitation; e.g. opponent drops price by 5 USD, a pure imitation would be an increase of 5 USD. $s = 0$ if agent-j has been acting in good faith throughout the negotiations, else $s = 1$.

I also apply a discount factor $0 < \gamma \leq 1$ to the value expected by a seller. γ, k, $V_{seller}(O)$ are the discount factor, current round k and the seller's evaluation of opponent offer O respectively. Such that:

$$\gamma^k V_{seller}(O) \tag{23}$$

So long as a buyer is acting cooperatively during a negotiation thread, a seller should steadily decrease the rigidity with which they evaluate their [Buyer] propositions. This is a useful concept that can capture advanced agreement structures like preferential treatment (i.e. reduced rates) for high-volume or loyal customers. I simplify the algorithm by having the discounting be affected only by the current round k with the assumption being that an opponent deserves leniency if it continues negotiations in good faith. The discounting factor can of course take into account more considerations, such as the weightings accorded to each issue by the seller. As to what constitutes a good attitude in a negotiating opponent, I define *negotiation momentum*, the value trajectory of an opponent's proposals.

Algorithm 2: Simple k-1 negotiation momentum

input : Negotiation thread N (list)
output: Boolean, True if opponent negotiating in good faith, else False

```
begin
    if len(N) ≤ 2 then /* Ensure previous history with opponent exists */
        break;
    end
    return V*_seller(N[k]) ≥ V*_seller(N[k − 1])
end
```

The algorithm above can be averaged over more steps.

3.3.4. Tactic Effectiveness. It is worth describing measures so that an agent, (here an abstracted buyer and seller), can evaluate the effectiveness of its tactics so that it may learn for future negotiations. *Intrinsic agent utility*. This is the measure which an agent would use to evaluate the utility of a negotiation's outcome. For simplicity, I consider price to be the only issue. This is a linear scoring function given agreed price x – when no deal is reached, the utility is 0.

$$U^{\text{buyer}}_{\text{price}}(x) = \frac{max^{\text{buyer}}_{\text{price}} - x}{max^{\text{buyer}}_{\text{price}} - min^{\text{buyer}}_{\text{price}}} \tag{24}$$

$$U^{\text{seller}}_{\text{price}}(x) = \frac{x - min^{\text{seller}}_{\text{price}}}{max^{\text{seller}}_{\text{price}} - min^{\text{seller}}_{\text{price}}} \tag{25}$$

Cost adjusted utility. It is worth knowing the relationship between an outcome's utility, the tactics ω used, and the costs involved – costs here could be direct (like taxes and other financial penalties) or more indirect (opportunity cost). We define some cost function, C, which contains all the resources an agent used during negotiations. Resource usage is further subject to a tax on communications. Each tactic's cost is unbounded, the cost need not be proportional to its percentage of

[26]Please pay attention to the superscripts here.

the agent's behavior. (Say imitation factors only as 5% of the agent's behavior but represents 30% of the resources used to reach settlement). $U^* = U - C(\omega)$.

Something worth considering is the informational content of a negotiation. On most issues, if no deal is reached, the agent would think the negotiation thread's utility is zero. However, even if no deal is reached, there ought to be a measure for the informational value of the negotiation (i.e., referrals and recommendations, market insights) – in line with our endeavor to make agents intelligent. It may be consistently true that an agent has a lot to learn from its opponent regarding the external landscape despite failed negotiations.

4. Attitudes and Policy

Reform is care after CA.

5. Conclusion

References

[1] Michael L. Littman *Markov games as a framework for multi-agent reinforcement learning*.

[2] Lutz Lampe, Mohamed O. Ahmed *Power Grid Topology Inference Using Power Line Communications*.

[3] Bruno Abrahao, Flavio Chierichetti, Robert Kleinber, Alessandro Panconesi *Trace Complexity of Network Inference*.

[4] Reid G. Smith *The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver*.

[5] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell *Overcoming catastrophic forgetting in neural networks*.

[6] David D'Achiardi, Nayara Aguiar, Stefanos Baros, Vijay Gupta and Anuradha M. Annaswamy *Reliability Contracts Between Renewable and Natural Gas Power Producers*.

[7] Nathanael Chambers, Dan Jurafsky *Unsupervised Learning of Narrative Event Chains*.

[8] Sarit Kraus, Katia Sycara, Amir Evenchik *Reaching agreements through argumentation: a logical model and implementation*.

[9] Iyad Rahwan *Interest-based Negotiation in Multi-Agent Systems*.

[10] Lamine Mili *Taxonomy of the Characteristics of Power System Operating States*

[11] T. Jaakkola and R. Barzilay *6.036 Spring 2018 : Week 10, RNN*.

[12] Stuart Russell *Rationality and Intelligence*.

[13] Le Xie, Dae-Hyun Choi, and Soummya Kar *Cooperative Distributed State Estimation: Local Observability Relaxed*.

[14] Baocheng Wang, Yafei Hu, Yu Xiao and Yi Li *An EV Charging Scheduling Mechanism Based on Price Negotiation*.