# Protégé: An Approach to Agental Collaboration

By D.M. Isom

**Abstract:**
Outlined below is an open framework for organizing agential cohorts against a broad range of applications. The framework scales well across various organizational sizes. It is dynamic enough as to allow its components to fit their environment with an appreciable degree of freedom. This would indicate that it should scale well across geographic distances.

The framework is intended to be populated by both ML learners (intelligence-on-silicon), as well as by human learners (intelligence-on-carbon). Because of this we will be opening with a brief section on the optimization of training for the human learner.

After this, we'll talk about *The Protégé Method*, the three phase training method of WATCH. DO. TEACH.

Then finally, we'll talk about the structure of the framework itself, with special emphasis on its method of distilling information down its layers, as well as how it reacts to novel situations.

*This is a companion piece to the paper "On the Psychology of Self-Organizing Machines."*

## The Structure of Learning

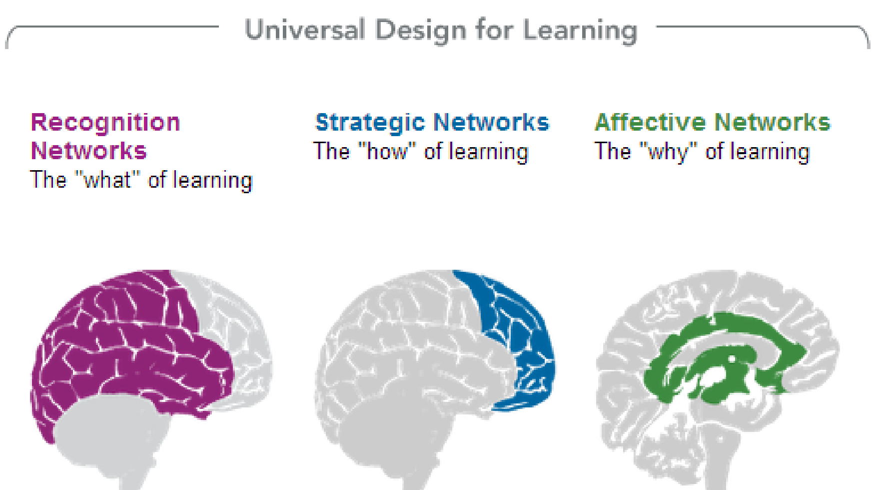
This section is meant to provide a high-level overview of the various systems used to inform the development of training regiment as well as a rudimentary understanding of the neuroscience at play.

This section is not meant to discuss the propagation of knowledge (which is outlined later), but rather a meta-overview on the construction of the material itself. First, we will briefly highlight the neural networks within our brains as they pertain to memory consolidation. Then, we'll discuss the various sensory modalities and their influence on our memory. Lastly, we'll go over the various patterns of learning and how they can be leveraged to improve overall retention.

### The Neurocognitive Structure of Learning

Not all learning is created equal. We are stimulus-driven, pattern detecting machines with no one set way of interpreting our observed world. The sights, the sounds, even the scents of things influence the formation of our memories and thus impact our learning. By understanding the different networks within our brains (and their role within our education) we can better leverage the neurocognitive sciences to make education more stimulating, engaging, and its effects longer lasting.

To start, let us get a basic understanding of some key regions within our brains and the cognitive networks therein.




The *Affective networks* (rightmost image) are responsible for our emotionally-laden experiences. To reference back to our "On the Psychology- " paper, this network is homologous to an ML learners *FLC subsystem*. This means that it is responsible for our motivational drive, reward function, and reinforcement-based learning.

To facilitate that reinforcement function, when we stimulate areas within these networks, our limbic system releases dopamine. Then when we pursue an interest (aka a goal), and should that pursuit result in a positive outcome, we then establish a positive feedback loop (ala reinforcement learning). Now the next time we're presented with that situation our limbic system will release even more dopamine and we'll be driven even harder in the pursuit of whatever that goal might be. However, should that initial interest not result in a positive outcome (even had the outcome not been particularly negative) then next time we will receive less dopamine—less motivation. This makes evolutionary sense, as why should an organism expend energy on something which nets it zero gains?

The *Recognition networks* (leftmost image) are where we process and store learned information. These networks are largely autonomic, representing the majority of an ML learner's *compute surface*. The principal component of this surface (at least where humans are concerned) is the cortical column, which itself is a sensory-motor prediction engine. These engines have become highly tuned over evolutionary time, and as such their resulting efficiency has become a bit of a two-edged sword.

On the one hand, the efficiency of the Recognition networks is essential (imagine the exhaustion you'd experience if everyday you took a new route into work). However, as helpful as this autonomic function may be, its efficiency can be problematic. Think of any time you've studied up until the point of boredom, for instance. You still recognize the pattern of words, the grammar, their structure, but you no longer understand their meaning. The task has lost its salience and become relegated to the unconscious, autonomic regions of your mind. Now the memories fade as quickly as they're formed.

The *Strategic network* (middle image). This is a *subsection* of the *compute surface* which has been partitioned off to the various components of the FLC system. Whereas most of the compute surface is dedicated to the processing of information as provided to it by various environmental sensors, this partition is instead allocated to the myriad subcortical structures found within the limbic system, and is thus truly general-purpose in its nature.

This brings with it several benefits, as well as some key disadvantages. The cortical column (as mentioned earlier) is largely convolutional in its processing function. This means the strategic networks are great for modeling complex structure (social concepts, problem spaces, visuospatial objects, et cetera), however that same convolutional architecture also means that the networks lack robust sequential recurrency. This is why it's difficult to remember more than seven numbers, in sequence, for instance. The emulation of such recurrent function is possible, albeit at great energetic cost.

By fully utilizing these three regions, and by understanding their functions and their uses, we can structure our trainings in such a way that not only are they more engaging, but more impactful too. Before we continue, let's recap each of the networks:

- **The Affective Networks**
    - Responsible for reward function, reinforcement learning, and motivation.
        - This means the Affective networks are goal-deterministic (They establish our wants).
    - Is capable of both positive-reinforcement as well as negative-reinforcement learning

- **The Recognition Networks**
    - The principal components of the Recognition networks find their function in the predictive processing of information.
        - This means they seek to recognize the underlying patterns hidden within datasets. Should these patterns be found, and accurate predictions therefore made, attentional awareness will localize upon the respective cortical column, leading to autonomic behavior.
            - Thus, if learning is to be conducted squarely through these networks then that learning ought be novel in its approach. Barring this, introducing stochasticity into the training regiment may help to refocus global attention.

- **The Strategic Networks**
    - The Strategic networks occupy the same computational surface as the Recognition networks.
        - Due to this these networks possess many of the convolutional advantages as seen by the rest of the compute surface, with a particular emphasis on a high modeling capacity.
        - However, due to this very same architecture, sequential memorization finds itself disadvantaged within this network.
            - Recurrent emulation *is* possible across convolutional space, however at great cost to the ML Learners attentional bandwidth.

## The Variability of Iconic, Echoic, and Haptic Memory

Modality. We've been tossing that word around for some time now, we should probably define what it means. A modality is simply a mode in which something is experienced or expressed. In the context of learning we're referring to "sensory modalities" (that is your five basic senses). The three we'll be referring to in this paper are iconic (sense of vision), echoic (sense of hearing), and haptic (sense of touch).

Each modality is expressed uniquely in the brain. For instance, your iconic memory has its own dedicated regional lobe (known as the occipital lobe) located in the rear of your skull.

We can calculate the specific data-throughput rates and persistence for each of the modal sources by taking the total bit-throughput rate of the human visual system (*~10,000,000 bits / per second*) and dividing that against the total number of cortical columns present within area V1 of that system. This nets us a rough estimate of approximately *650 bits / per second / per column*.

We can then take this value and divide against the known bit-throughput rates of the echoic and haptic systems. For hearing, that rate is *100,000 bits / per second* - which gives us *~150 columns* present in area A1. As for the haptic system, we process *1,000,000 bits / per second*. This grants us *~1,500 columns* as present in area S1 of the somatosensory cortex.

These measurements are important because they establish for us the total width of each of the modalities input layers. A larger width grants us a richer sampling rate of the world (aka its of a higher resolution).

We can then determine the persistence of that throughput by calculating its synaptic transmission route found between its time of input and its archival within the hippocampal formation. Shorter routes equate to higher persistence.

Termination into the data archival system is important because the trisynaptic circuit, unlike its overlying compute surface, possesses true sequential recurrency. This to say, it remembers.

Obviously, learning is more nuanced than simply memorizing throughputs and persistence. For instance, in spite of the 11,000,000 bit rate found on the compute surface, the attentional guidance system (AGS) only samples 40 - 60 bits of that per second (this being the actual inferenced memory you're aware of at time of operation). Then there are also other important metrics; recurrent encoding within the hippocampus, for instance– where long-term memory storage encodes at theta-rhythm (8-12Hz). Then there's the interplay between all these systems, and their compound effects on learning. For example, that aforementioned theta-rhythm is modulated by a structure known as the medial septum, which itself is activated by haptic activity. So yes, movement has a role to play in learning.

## A Note on Fatigue

Exhaustion is the buildup of debris as deposited within and between the synaptic clefts of neurons. This debris then obscures that synaptic transmission space, "muffling" the neuromodulatory communications taking place there. It is this "fuzzing" of the signal that leads to the various phenomena expressed whilst one is operating within a state of fatigue.

If developing embodied systems which leverage any form of extracellular fluidity (which *does* have its advantages found in its efficiency), just know that you will likely need to contend with similar constraints. Plan accordingly.

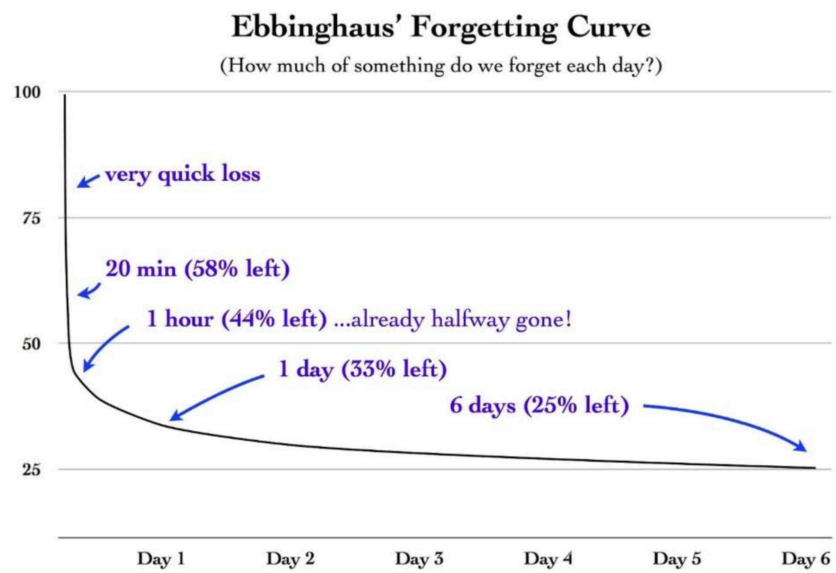
## Interleaving Data Sets

We–as with all intelligent machines–are prone to both the overfitting and underfitting of learned information. The former happens when we train specifically on an exclusive data set. Just as with an ML learner we are *only* memorizing the data, and we are not learning the underlying algorithms and patterns hidden within it. Inversely, the latter occurs when we train on over complex datasets without having first learned its constituent parts (remember, the human's compute surface is convolutional in its architecture. 'Simple things into complex things').

An optimal approach is to instead train on appropriate stochastic data sets which, when sufficiently generalized, find commonality with one another. Doing this will leverage the passive benefits of an ML learners Memory-tuning subsystem (which in humans is the *Default Mode Network*). Doing this will facilitate learning long past time-of-training, while the agent is at rest.

### The Recall Advantage

Well-studied research suggests that we forget roughly 70% of what we learn within one day having learned it.




The reasons for this are several fold:

(1) It *is* possible that Memory-tuning is over-generalizing and thus reducing information to the point of irrecoverability. However, the graph shows a 65% loss rate within an hour post-training. This generally isn't enough time for the Default Mode Network to have traced out sufficient manifolds within high-dimensional space. Instead, the culprit is likely item number two…

(2) Earlier we had discussed simulating recursion while operating out from convolutional architecture. This emulation occurs within the Strategic networks (that partitioned section of the compute surface). Do you recall how I mentioned that doing this comes at a great cost? I suspect that this graph *is* that cost. In a sense, we are exceeding our biological context windows.

So what are some strategies to negate this loss curve?

**Overlearning vs. Earlystopping**

*Overlearning* refers to the training on information even after the point of initial mastery. A simple example of this would be like the re-studying of fundamentals of a field for which you're already proficient.This is essentially overfitting by design, and is often used in conjunction with training techniques such as minibatching. A greater subdivisioning of the total batch typically leads to superior, albeit artificial, gains in the near-term.

Studies show that participants trained on overlearning indeed had superior memory retention in days immediately following training. However, in the weeks and months after, that retention decayed off to the point of becoming negligible. It should be stated that the time-frames observed post-decay are more than long enough for Memory-tuning to have regularized the data into irrelevancy.

*Earlystopping* employs many of the same ideas as overlearning, however, its goal is to cut off training at the point which the learning begins to overfit; its focus being on optimizing for quick generalizations of the data. Earlystopping is susceptible to many of the same issues as overlearning (where Memory-tuning is concerned).

**The Spacing Effect**

The spacing effect introduces some slight changes to the typical process of mini-batched training. With it, the ML learner paused between training epochs. When the next training cycle is set to begin, the trainer is asked to recall the final moments of the previous epoch. We may view this as a kind of self-reflection of the training data. You can find current research on the spacing effect here.

*This would be an area benefited by further study…*

## Putting it all Together

The purpose of all this is to highlight some of the cortical functions as they pertain to our education. Its important that we understand the strengths and weaknesses found within each mode of learning. This chapter should help inform us when it comes to the design of online courses, but it could also prove useful in providing live training.

The only mechanism for learning we haven't covered here yet is that of social cognition. Perhaps the most important mechanism, we'll be spending the remainder of the paper discussing how to best leverage it. That said, the next chapter will highlight a teaching methodology called *The Protégé Method*, which takes into consideration nearly all of what we've discussed and applies it to a near-peer education system. Until then, here is a recap of this chapter.

- There are **three cognitive networks** in our brains (as they pertain to education):
    - Affective**:** The emotional region of our brain. Responsible for creating the "initial interest". Stimulation of this region is required for focused attention. Can create feedback loops which can be positive or negative.
    - Recognition**:** An region of the brain whose principal components seek to identify patterns within the provided data. Should their predictive processing be met with relative success, attention will localize upon the unit, thereby becoming autonomic.
    - Strategic**:** The "conscious thought" region of our brain. Not conscious as in 'consciousness,' but rather requiring active participation. Keeping this region engaged is essential for learning new material. Also, the most energetically demanding region–requiring frequent breaks.
- **Variability in Memory**. (1) Iconic memory is high bandwidth, low persistence. Great for ingesting high-fidelity, albeit short-lived, material. (2) Echoic memory is low bandwidth, high persistence. Great for forming long-term memories, though this encoding process takes time. (3) Haptic memory sits somewhere between the other two and is more variable. Its greatest strength may actually lie in its forced modulation of theta-rhythm in long-term archival.
- **Interleaving Data Sets.** This works the brain's ability at pattern recognition. If varying sensory modalities stimulates the Affective Networks, this does the same for the Strategic networks. Additionally, by mixing associative functions we can more easily highlight their interconnections.
- **Ebbinghaus' Forgetting Curve**. Within twenty minutes we forget roughly half of what we've learned. Consciously changing our patterns of learning and encouraging mnemonic recall helps to improve this.
- **The Spacing Effect.** This helps to reset the Ebbinghaus Curve. Spacing out testing and discussion forces us to recognize associated patterns in learned material, further strengthening those memories.

## The Protégé Method

The three-phase training cycle of WATCH. DO. TEACH.

The Protégé method takes influence from the *protégé effect*. This psychological effect is a well-documented phenomenon in which the act of teaching serves as a mode for learning, not just for the student, but for the mentor as well.

The Protégé Method follows three basic rules:

1. **WATCH**
   The protégé(s) are taught by their mentor on specific knowledge set(s). The mode by which training is administered is left open to interpretation, with full creative freedom given over to the mentor to teach the subject as they see fit (i.e. they're not forced to teach it in any one specific way).

2. **DO**
   Immediately following the WATCH phase (or shortly after) the protégé(s) are enrolled in a structured knowledge-base course in which they relearn the data as recently taught to them by their mentor. It is at this stage that a governing body (for example, a SOB) may establish a weighted standard by which the collective whole is held to. At the end of this exercise the protégé(s) will be made to take an evaluation. This evaluation doesn't grade the material itself, but rather it serves as a critique of the previous mentor's delivery of that material (i.e., how well the mentor trained on it). This serves two functions:

   a. To provide feedback to the mentor. In doing this they can more easily identify gaps in their own personal knowledge. This also serves to provide an underlying motivational function to the mentor agent.

   b. To provide analytic data to the governing body. If broad trends are noticed in how certain things are being trained, they can adjust for that in their updates to the global knowledge-base.

      ***To discourage the mentor from overfitting upon the global standards dataset's, I would suggest implementing some form of gradient-based negative reward function. With this you may be able to establish an operational band, inside of which creative interpretation of the global data may be expressed in agental training.*

3. **TEACH**
   The final phase is for our protégé(s) to become mentors themselves. They will be tasked with disseminating the training down to the next layer (should one exist) knowing that they too will be evaluated by their own protégé(s).

**Distillation & Response**

These next two chapters will discuss the propagation of training. This first chapter will cover higher-level concepts, whereas the second chapter will be more anecdotal in nature.

### The Nature of Top-Down Structure

The nature of hierarchical distribution is that every rung of the structure has the basic working knowledge of each rung that preceded it. In theory, as one climbs the hierarchical structure the breadth of their knowledge widens, however perhaps becoming lower in fidelity. Inversely, as one descends the structure the breadth of their understanding narrows but becomes of a higher resolution–often fitting to their localized environment.

The example we will be using for our demonstration will be that of an agential cohort specced for property management. However, the concepts here are flexible, and the same ideas may be applied to anything from the training of software development teams, to the fitting of autonomous fighter-squadrons.

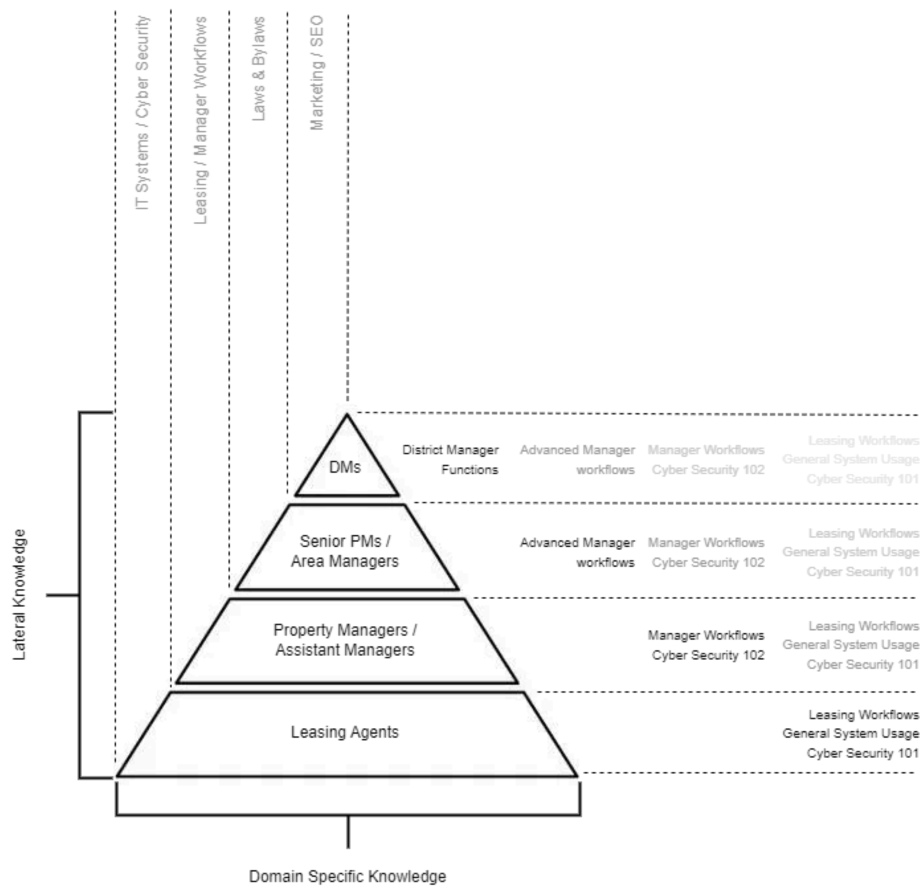



*Figure 2*

*An abstraction of top-down structure. Notice how the lateral knowledge is an aggregate across all domain spaces, whereas domain specific knowledge is a narrow band of highly specialized information. In this way, we can view lateral knowledge as a shallow ocean, and domain-specific knowledge as a deep pond.*

In addition to this there exist domain-specific specializations. These positions consist of highly specialized individuals (or groups thereof) who demonstrate both a fluency and a mastery at every echelon within the structure, albeit operating within a narrow band. These specialists may form the regulatory arm of the governing body, or they may be of the governing body itself.

Distillation of knowledge should be that the domain-specific roles work in collaboration with the uppermost rung of the agental cohort, wherein the flow of that knowledge is propagated ever downwards. In context to our property management example, the Specialist would educate the District Managers > who would then educate their respective Senior Managers > who then educated their PMs > who then educate their leasing agents. As this distalls down

If you choose to incorporate this convolutional approach, please note the following: it is critical that you not allow knowledge to gap rungs. In doing this we allow for levels of plausible deniability to exist. That is, there would exist individuals who would be exempt from the standards potentially expected of their subordinates. Furthermore, these individuals would likely be the ones responsible for upholding those standards for which they themselves are not measured against. Standard should rise with station.

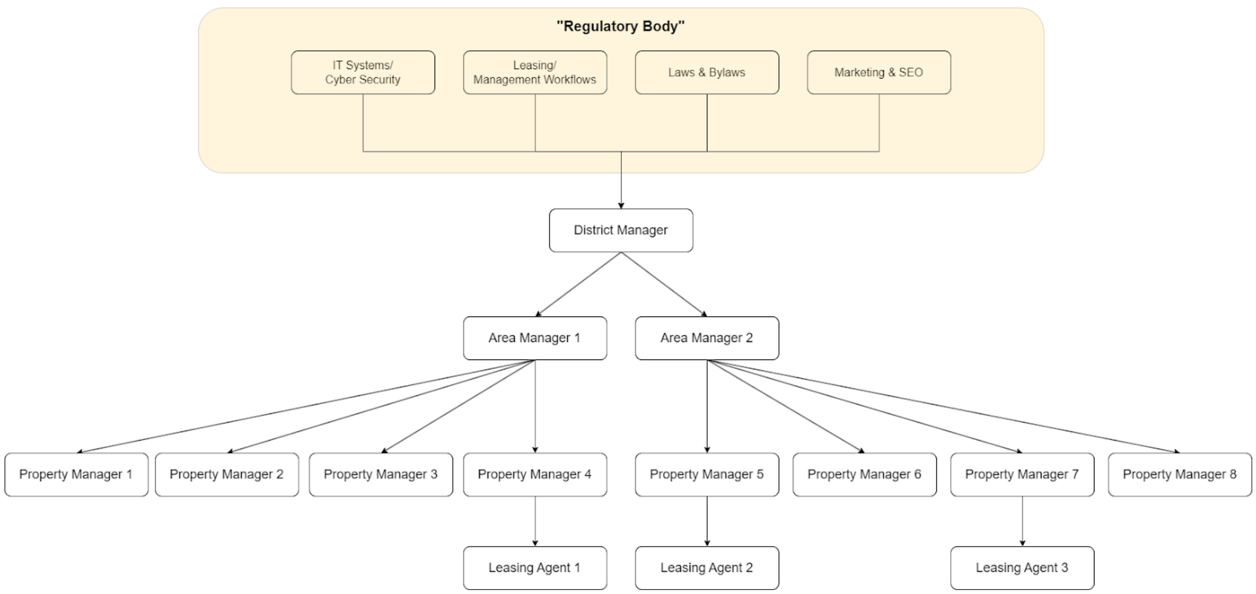



*Figure 3*

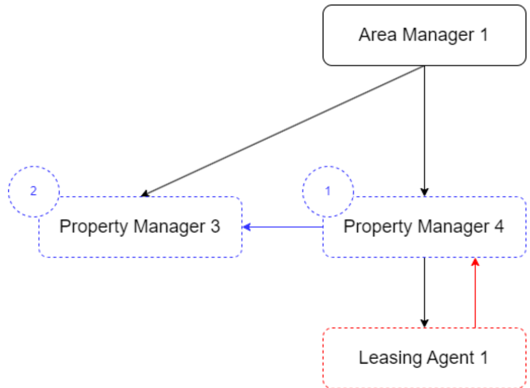
A technical. In this case, the Specialists would provide support and training to the District Manager, who would then train their Area Managers, who would then continue providing the training downwards. Each level is accountable for the training of the level directly beneath them.

### The Rule of Over & Up

This approach of *distillation* and *response* carries with it two key advantages. The first, as outlined above, is in its efficient centralized structure of passing information down throughout a decentralized apparatus, whilst allowing for maximal adaptability at its lowermost components.

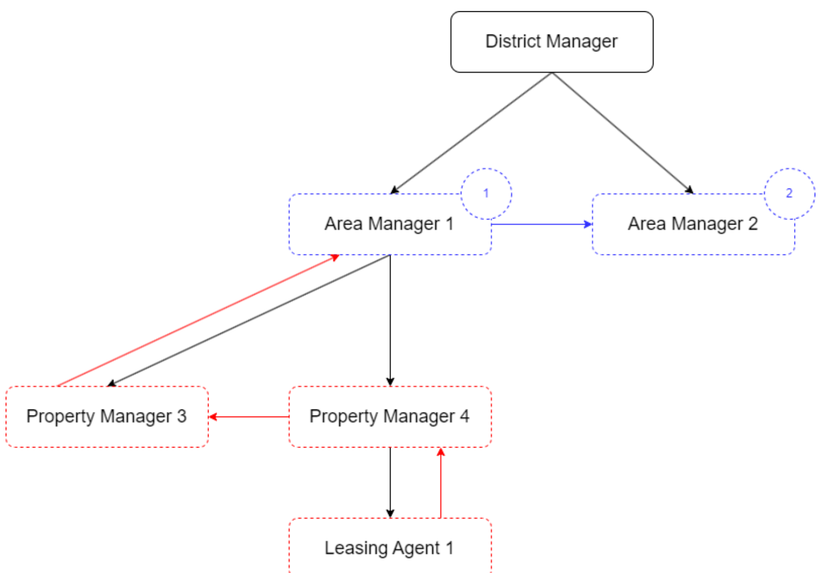
The second advantage is found in the autonomy of those lower regions. One of the main disadvantages in many centralized structures is the time it takes for lower-level nodes to send and receive communications from higher-order administrative bodies. The approach listed below would seek to remedy this. It would push all relevant information as close to *surface level* as possible (surface level being the nodes near or at the bottom).

Of course, novel situations may still arise to which local level nodes have no generalized response (situations referred to as “Outside Context Problem”). However, the structure of the system itself adds an additional layer of intelligence, independent of the agents. This extra layer makes for an adaptable structure that is capable of responding to situations in dynamical ways, learning over time. Let's explore how it works with a simple example:



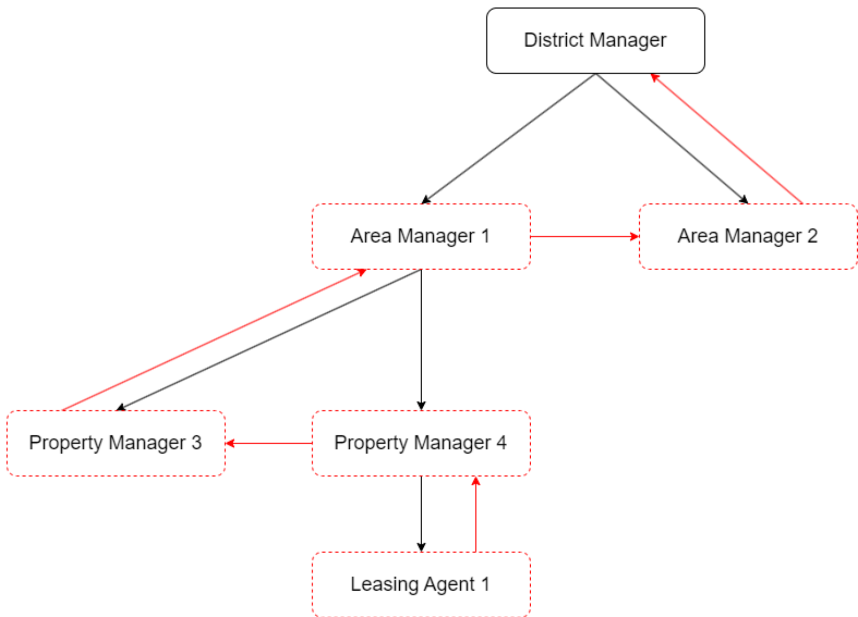

*Stage 1*

*In this instance we'll pretend that the leasing agent "Leasing Agent 1" doesn't know how to perform a certain leasing function. Because the agent is the only leasing agent at this specific property they would reach* up *to their manager, "Property Manager 4." For now, let's pretend that this manager also doesn't know how to perform the function. They then reach* over *to the manager at a neighboring property. Now, if Property Manager 3 also does not know the answer, then we have identified a problem in the propagation of that training. In this example it's likely that training is not being disseminated down from Area Manager 1.*




*Stage 2*

We can see how, when a lack of knowledge presents itself, an agent should bring it to the attention of another agent at their level. Then, if that second person should also lack the answer, they then raise the problem up to the level above them. In this case Property Manager 4 would bring the question up to Area Manager 1. The process repeats until an answer is found. This is the rule of *Over and Up*.




*Stage 3*

The *Over and Up* rule leads to what I refer to as the "cascade effect." Its purpose is to create a slow but mounting pressure to address gaps in knowledge. However, instead of cascading down it cascades up. In doing it this way we can compartmentalize deficiencies to the smallest common denominator and, in theory, the cascade stops where the problem began. The probability distribution behind this is simple. The likelihood of two agents being taught something, yet neither of them being able to remember it, is negligibly low. However, the likelihood of neither being taught it to begin with is high. Moreover, if they've not been trained then the likelihood of others agents within their local area *also* not knowing is likewise very high.

## Key Concepts and Takeaways

1. **The Protégé Method**
   *The three-phase training cycle of Watch. Do. Teach.*

   This training method takes advantage of nearly every cortical function within the brain to ensure greater knowledge retention. The phases consist of (1) a mentor teaching a subject to their proteges, (2) those proteges then taking an centralized course which pertains to that subject at the end of which they critique the previous mentors training, and (3) the proteges then become the mentors themselves. The cycle repeats.

2. **Top-Down Training**
   *Training and support auto-distributes and load-balances.*

3. **The Rule of Over & Up**
   *We compartmentalize issues to their lowest common denominator, and allow for fast local response times.*