Babble and Prune

1. Babble
2. More Babble
3. Prune
4. Circumambulation
5. Write

Babble

This is a linkpost for https://radimentary.wordpress.com/2018/01/10/babble/

This post is an exercise in "identifying with the algorithm." I'm a big fan of the probabilistic method and randomized algorithms, so my biases will show.

How do human beings produce knowledge? When we describe rational thought processes, we tend to think of them as essentially deterministic, deliberate, and algorithmic. After some self-examination, however, I've come to think that my process is closer to babbling many random strings and later filtering by a heuristic. I think verbally, and my process for generating knowledge is virtually indistinguishable from my process for generating speech, and also quite similar to my process for generating writing.

Here's a simplistic model of how this works. I try to build a coherent sentence. At each step, to pick the next word, I randomly generate words in the category (correct part of speech, relevance) and sound them out one by one to see which continues the sentence most coherently. So, instead of deliberately and carefully generating sentences in one go, the algorithm is something like:

1. Babble. Use a weak and local filter to randomly generate a lot of possibilities. Is the word the right part of speech? Does it lie in the same region of thingspace? Does it fit the context?
2. Prune. Use a strong and global filter to test for the best, or at least a satisfactory, choice. With this word in the blank, do I actually believe this sentence? Does the word have the right connotations? Does the whole thought read smoothly?

This is a babble about embracing randomness.

Baby Babble

Research on language development suggests that baby babble is an direct forerunner to language. You might imagine that infants learn by imitation, and that baby babble is just an imperfect imitation of words the baby hears, and progress occurs as they physiologically adapt to better produce those sounds. You would be wrong.

Instead, infants are initially capable of producing *all the phonemes* that exist in all human languages, and they slowly prune out which ones they need via reinforcement learning. Based on the sounds that their parents produce and respond to, babies slowly filter out unnecessary phonemes. Their babbles begin to drift as they prune out more and more phonemes, and they start to combine syllables into proto-words. Babble is the process of generating random sounds, and looking for clues about which ones are useful. Something something reinforcement learning partially observable Markov decision process I'm in over my head.

So, we've learned that babies use the Babble and Prune algorithm to learn language. But this is quite a general algorithm, and evolution is a conservative force. It stands to reason that human beings might learn other things by a similar algorithm. I don't think it's a particularly controversial suggestion that human thought proceeds roughly by

cheaply constructing a lot of low-resolution hypotheses and then sieving from them by allowing them to play out to their logical conclusions.

The point I want to emphasize is that the algorithm has two distinct phases, both of which can be independently optimized. The stricter and stronger your Prune filter, the higher quality content you stand to produce. But one common bug is related to this: if the quality of your Babble is much lower than that of your Prune, you may end up with nothing to say. Everything you can imagine saying or writing sounds cringey or content-free. Ten minutes after the conversation moves on from that topic, your Babble generator finally returns that witty comeback you were looking for. You'll probably spend your entire evening waiting for an opportunity to force it back in.

Your pseudorandom Babble generator can also be optimized, and in two different ways. On the one hand, you can improve the weak filter you're using, to increase the probability of generating higher-quality thoughts. The other way is one of the things named "creativity": you can try to eliminate systematic biases in the Babble generator, with the effect of hitting a more uniform subset of relevant concept-space. Exercises that might help include expanding your vocabulary, reading outside your comfort zone, and engaging in the subtle art of nonstandard sentence construction.

Poetry is Babble Study

Poetry is at its heart an isolation exercise for your Babble generator. When creating poetry, you replace your complex, inarticulate, and highly optimized Prune filter with a simple, explicit, and weird one that you're not attached to. Instead of picking words that maximize meaning, relevance, or social signals, you pick words with the right number of syllables that rhyme correctly and follow the right meter.

Now, with the Prune filter simplified and fixed, all the attention is placed on the Babble. What does it feel like to write a poem (not one of those free-form modern ones)? Probably most of your effort is spent Babbling almost-words that fit the meter and rhyme scheme. If you're anything like me, it feels almost exactly like playing a game of Scrabble, fitting letters and syllables onto a board by trial and error. Scrabble is just like poetry: it's all about being good at Babble. And no, I graciously decline to write poetry in public, even though Scrabble does conveniently rhyme with Babble.

Puns and word games are Babble. You'll notice that when you Babble, each new word isn't at all independent from its predecessors. Instead, Babble is more like initiating a random walk in your dictionary, one letter or syllable or inferential step at a time. That's why word ladders are so appealing - because they stem from a natural cognitive algorithm. I think Scott Alexander's writing quality is great partly because of his love of puns, a sure sign he has a great Babble generator.

If poetry and puns are phonetic Babble, then "Deep Wisdom" is semantic Babble. Instead of randomly arranging words by sound, we're arranging a rather small set of words to sound wise. More often than not, "deep wisdom" boils down to word games anyway, e.g. wise old sayings:

"A blind person who sees is better than a seeing person who is blind."

"A proverb is a short sentence based on long experience."

"Economy is the wealth of the poor and the wisdom of the rich."

Reading is Outsourcing Babble

Reading and conversation outsource Babble to others. Instead of using your own Babble generator, you flood your brain with other people's words, and then apply your Prune filter. Because others have already Pruned once, the input is particularly high-quality Babble, and you reap particularly beautiful fruit. How many times have you read a thousand-page book, only to fixate on a handful of striking lines or passages?

Prune goes into overdrive when you outsource Babble. A bug I mentioned earlier is having way too strict of a Prune filter, compared to the quality of your Babble. This occurs particularly to people who read and listen much more than they write or speak. When they finally trudge into the attic and turn on that dusty old Babble generator, it doesn't produce thoughts nearly as coherent, witty, or wise as their hyper-developed Prune filter is used to processing.

Impose Babble tariffs. Your conversation will never be as dry and smart as something from a sitcom. If you can't think of anything to say, relax your Prune filter at least temporarily, so that your Babble generator can catch up. Everyone starts somewhere - Babbling platitudes is better than being silent altogether.

Conversely, some people have no filter, and these are exactly the kind of people who don't read or listen enough. If all your Babble goes directly to your mouth, you need to install a better Prune filter. Impose export tariffs.

The reason the Postmodernism Generator is so fun to read is because computers are now capable of producing great Babble. Reading poetry and randomly generated postmodernism, talking to chatbots, these activities all amount to frolicking in the uncanny valley between Babble and the Pruned.

Tower of Babble

A wise man once said, "Do not build Towers out of Babble. You wouldn't build one out of Pizza, would you?"

NP

NP is the God of Babble. His law is: humans will always be much better at verifying wisdom than producing it. Therefore, go forth and Babble! After all, how did Shakespeare write his famous plays, except by randomly pressing keys on a keyboard?

NP has a little brother called P. The law of P is: never try things you don't understand completely. Randomly thrashing around will get you nowhere.

P believes himself to be a God, an equal to his brother. He is not.

More Babble

This is a linkpost for https://radimentary.wordpress.com/2018/01/11/more-babble/

In my last babble, I introduced the Babble and Prune model of thought generation: Babble with a weak heuristic to generate many more possibilities than necessary, Prune with a strong heuristic to find a best, or the satisfactory one. I want to zoom in on this model. If the last babble was colored by my biases as a probabilist, this one is motivated by my biases as a graph theorist.

First, I will speculate on the exact mechanism of Babble, and also highlight the fact Babble and Prune are independent systems that can be mocked out for unit testing.

Second, I will lather on some metaphors about the adversarial nature of Babble and Prune. Two people have independently mentioned Generative Adversarial Networks to me, a model of unsupervised learning involving two neural nets, Generator and Discriminator. The Artist and the Critic are archetypes of the same flavor - I have argued in the past the spirit of the Critic is Satan.

Babble is (Sampling From) PageRank

Previously, I suggested that a Babble generator is a pseudorandom word generator, weighted with a weak, local filter. This is roughly true, but spectacular fails one of the technical goals of a pseudorandom generator: independence. In particular, the next word you Babble is frequently a variation (phonetically or semantically) of the previous one.

PageRank, as far as I know, ranks web pages by the heuristic of "what is the probability of ending up at this page after a random walk with random restarts." That's why a better analogy for Babble is sampling from PageRank i.e. taking a weighted random walk in your Babble graph with random restarts. Jackson Pollock is visual Babble.

Imagine you're playing a game of Scrabble, and you have the seven letters JRKAXN. What does your algorithm feel like?

You scan the board and see an open M. You start Babbling letter combinations that might start with M: MAJR, MRAJ, MRAN, MARN, MARX (oops, proper noun), MARK (great!). That's the weighted random walk. You set MARK aside and look for another place to start.

Time for a restart. You find an open A before a Triple Word, that'd be great to get! You start Babbling combinations that end with A: NARA, NAXRA, JARA, JAKA, RAKA. No luck.

Maybe the A should be in the middle of the word! ARAN, AKAN, AKAR, AJAR (great!). You sense mean stares for taking so long, so you turn off the Babble and score AJAR for (1+8+1+1)x3 = 33 points. Not too shabby.

The Babble Graph

Last time, I described getting better at Babble as increasing the uniformity of your pseudorandom Babble generator. With a higher-resolution model of Babble in hand, we should reconceptualize increasing uniformity as building a well-connected Babble graph.

What is the Babble graph? It's the graph within which your words and concepts are connected. Some of these connections are by rhyme and visual similarity, others are semantic or personal. Blood and snow are connected in my Babble graph, for example, because in Chinese they are homophones: snow is 雪 (xue), and blood is 血 (xue). This led to the following paragraph from one of my high school essays (paraphrased):

> In Chinese, snow and blood sound the same: "xue." Some people think the world will end suddenly in nuclear holocaust, pandemic, or a belligerent SkyNet. I think the world will die slowly and painfully, bleeding to death one drop at a time with each New England winter.

My parents had recently dragged me out to jog in the melting post-blizzard slush.

One of my favorite classes in college was a game theory class taught by the wonderful David Parkes; my wife and I lovingly remember the class as Parkes and Rec. One of the striking ideas I learned in Parkes and Rec is that exponentially large graphs can be compactly represented implicitly in memory, as long as individual edges and neighborhoods can be computed in reasonable time. Babble is capable of generating new words and combinations, so the Babble graph contains nodes you've never thought of. It's enormous, and definitely not (a subgraph of) the connectome, but rather implicitly represented therein in a compact way. This is related to the fact that the map is not the territory, except in the study of the brain, where the map is a subset of the territory.

It follows that the Babble graph is a massive implicitly represented graph, which is traversed via random walks with random restarts. How might we optimize this data structure to better fulfill its goals?

One technique I've already mentioned is to artificially replace either Babble or Prune to train the other in isolation. This is basically unit testing via mocking. To unit test Babble, we can mock out Prune with a simplified and explicit filter like the haiku, the game of Scrabble, or word games like Contact and Convergence. To unit test Prune, we replace Babble with other sources of word strings: reading, conversation, poetry, music.

Those methods completely black box the Babble and Prune algorithms and hope they self-optimize correctly. What if we want to get our hands dirty and explicitly rewire our Babble graph?

First we have to figure out what makes a quality Babble graph. I can think of two metrics worth optimizing:

1. Connectivity. With sufficient effort (i.e. taking enough random steps and restarts) you want to eventually explore the entire graph, and repeat yourself rarely. This requires not just that the graph is connected, but that it should have good expansion. Ever feel stuck on an idea, then be struck by external inspiration to explore a disconnected set of ideas you already knew, and find it massively

productive? Random walks getting trapped locally is a sign your Babble graph is a bad expander.
2. Value. Every node in your Babble graph should pay rent. I have found many abandoned components in my Babble graph - ghost towns and wastelands of neural machinery left over from experiences that are no longer relevant. They can be salvaged and repurposed, if only to generate metaphors.

Ramanujan was an extraordinarily creative mathematician who produced formulas like

$$p(n) \sim \frac{1}{4n\sqrt{3}} \exp \pi \sqrt{\frac{2n}{3}}$$

for the number of partitions p(n) of an integer n. Exercise: figure out how such an exponent might occur in nature. Hint: $\zeta(2) = \pi^2/6$.

Ramanujan was also known for his mysticism, attributing his most inspired results to his patron goddess. Mystical experiences, like LSD, are often characterized by the feeling of connectedness of all things. I think Ramanujan's genius might be the result of having a Babble graph that is an exceptionally good expander. What are those called again?

Here's a story about how I improved my Babble graph by making my bed.

It all started when Jordan Peterson told me to clean my room - because one's surroundings are a reflection of one's state of being. I decided to give it a chance and make my bed every morning.

Making my bed became a daily ritual. As I do it, I repeat the "proper and humble" mantra:

To save the world, I will start by doing the proper and humble things I know how to do within the confines of my own life.

Proper and humble were not words I'd liked in a very long time. They activated ideas I haven't wrestled with for years.

Honte is a Go term which means "the proper move." Honte is playing thickly to leave few weaknesses. Honte is killing already dead stones to remove aji. Honte is doing the proper and humble thing to prevent bad aji - failure modes you can't yet articulate. There's nothing quite like playing Go against a stronger player to put the fear of aji in you.

In relationships, honte is dedication to the removal of lingering resentment. Unhappy couples have the same fights at regular intervals; the landmines that trigger them might lay untouched for upwards of a year, but they never deactivate. Why would you allow these landmines be planted in the first place? You wouldn't leave a ladder breaker for your opponent in an unapproached corner, would you? Dedicate yourself to the removal of landmines, at least when you have the slack to do so. That's honte.

A well-connected and useful Babble graph is thickness (not to be confused with thiccness). It is written: *attack from thickness*. When thinking from a thick Babble

graph, you're not wandering lackadaisically, building an argument from scraps lying at the side of the trail. You'll have the weight of your entire intellectual life at your back.

The Artist and the Critic

Two people have independently suggested that the Babble and Prune model is similar to an approach in machine learning known as Generative Adversarial Networks, in which production of photorealistic images (say) is turned into a game between two neural nets, Generator, who learns to generate good counterfeits, and Discriminator, who works on finding the real stuff.

This is a manifestation of the eternal war between the Artist and the Critic, a war that is both exceedingly vicious and exceedingly productive. Artists of the ages have had some choice words for their critics. Beckett:

> VLADIMIR
> Moron!
> ESTRAGON
> That's the idea, let's abuse each other.
> *They turn, move apart, turn again and face each other.*
> VLADIMIR
> Moron!
> ESTRAGON
> Vermin!
> VLADIMIR
> Abortion!
> ESTRAGON
> Morpion!
> VLADIMIR
> Sewer-rat!
> ESTRAGON
> Curate!
> VLADIMIR
> Cretin!
> ESTRAGON
> *(with finality)* Crritic!
> VLADIMIR
> Oh!
> *He wilts, vanquished, and turns away.*

The opening lines of Hardy's *A Mathematician's Apology*:

> It is a melancholy experience for a professional mathematician to find himself writing about mathematics. The function of a mathematician is to do something, to prove new theorems, to add to mathematics, and not to talk about what he or other mathematicians have done. Statesmen despise publicists, painters despise art-critics, and physiologists, physicists, or mathematicians have usually similar feelings: there is no scorn more profound, or on the whole more justifiable, than that of the men who make for the men who explain. Exposition, criticism, appreciation, is work for second-rate minds.

I have a rule inspired by Solzhenitsyn, which is that every battle which occurs between human beings also plays out within each human heart. The proper locus of the fight between Artist and Critic is not cleanly between artists and critics, but between the Babble and Prune within each individual. After all, find me an artist who has never criticized, or a great critic who is never enjoyable to read for his own sake. Exercise: get some utility out of a bad book you've recently read by checking out the savage reviews online.

Like Generator and Discriminator, a good Artist and Critic pair can together ascend to heights that neither could reach alone, and having a filter is a healthy thing. However, I stand by my argument that the overdeveloped Critic is a manifestation of Satan:

> Jordan Peterson says Satan is an intellectual figure, and this idea has fermented in my imagination. Satan is the cynical and nihilistic intellectual whose thesis is "things are so bad they do not deserve to exist."
> [...]
> I would propose an embellishment of the figure of Satan as the nihilistic intellectual: Satan as the critic. One of the (many) disturbing things I have noticed about my high school curriculum is that English classes are factories for creating critics out of artists. At least in my experience, we wrote short stories, poems, and other free form essays in elementary and middle school, but turned exclusively to the analytical essay by the time high school rolled around.
> How frightening is that? Take a generation of teenagers, present them with the greatest literature of our civilization. Then, instead of teaching them to do the obvious thing – imitate – we teach them to analyze – the derivative work of a critic. The work of Satan: the intellectual whose ability to criticize far exceeds his ability to create. And so we find that the best students to come out of our high schools are created in the image of Satan. For every one budding novelist, we have a dozen teenage journalists, lawyers, and activists.
> Satan is the voice in your ear who says, "You will never do this well enough for it to be worth doing." This is the burrowing anxiety that puts me off writing for weeks at at time, the anxiety that anything I produce will not justify its own existence. The subroutine in your head constantly constructing impossibly high standards and handing them to you to use as excuses to do nothing. Satan is characterized by inaction, the inaction caused by paralyzing perfectionism.

Other Things that are Babble

The Bible is the best Babble ever produced. A common atheist refrain is that the Bible is so self-contradictory, so ambiguous, so open to interpretation as to be intrinsically meaningless. Any meaning you might extract from the Bible is just a reflection of your own beliefs.

I think this is a feature, not a bug.

Not only is the Bible open to interpretation, it *invites* interpretation. Its stories are so varied, fantastical and morally ambiguous that they *demand* interpretation. The Bible stood the test of time not because it is maximally packed with wisdom, but because it produced the most insightful and varied results when paired with outside sources of Prune. When the Christian is lost and desperate, he inputs 1 Corinthians to his Prune, and voilà! Faith is restored. Peterson's The Psychological Significance of the Bible series takes advantage of exactly this feature of the Bible: it is the fertile ground

upon which each individual can tell their own story. Of course, perversions can result when broken Prune filters are applied, even to the best Babble.

Perhaps writers have been optimizing for the wrong thing. Instead of directly packing insight into an essay, we should try to design high-quality Babble, fertile input for the reader's Prune.

The Oulipo is Babble training on steroids - a group of writers and mathematicians who worked based on the apparent paradox that freedom is the enemy of creativity. Creativity, the state of having a better Babble generator, is designed to solve tough, heavily constrained problems, and the Oulipians produced creative writing by imposing stricter restraints. Most famously, this method produced Perec's novel *La disparition*, a 300-page novel written without the letter 'e,' about "a group of individuals looking for a missing companion, Anton Vowl."

By the way, did you notice the letter missing from this entire post?

All good conversations are therapeutic, and therapeutic conversations are about letting down your guard and allowing yourself to simply Babble. Babies have no Prune at all and babble all the phonemes their adorable little mouths can produce - that's how they learn the beginnings of language so quickly. Being in a safe space is reproducing this state of development, a place where Babble can be rapidly be optimized on its own terms. Healthy teamwork and collaboration shares this quality: bouncing half-formed, half-nonsensical ideas off others and Pruning them together. Double the Babble, double the fun.

Oh, and about that missing letter? Just kidding. Ain't nobody got time for that.

Prune

This is a linkpost for https://radimentary.wordpress.com/2018/01/12/prune/

Previously, I described human thought-generation as an adversarial process between a low-quality pseudorandom Babble generator and a high-quality Prune filter, roughly analogous to the Generative Adversarial Networks model in machine learning. I then elaborated on this model by reconceptualizing Babble as a random walk with random restarts on an implicitly stored Babble graph.

Rationalist training (and schooling in general) slants towards developing Prune over Babble. I'm trying to solve the dual problem: that of improving the quality of your Babble.

Although the previous posts listed a number of exotic isolation exercises for Babble, I'm guessing nobody was inspired to go out and play more Scrabble, write haikus, or stop using the letter 'e'. That's probably for the best - taking these exercises too seriously would produce exotic but sub-optimal Babble anyway. For a serious solution to this serious problem, we need to understand Prune at a higher resolution.

The main problem with Prune is that it has too many layers. There's a filter for subconscious thoughts to become conscious, another for it to become spoken word, another for the spoken word to be written down, and a further one for the written word to be displayed in public. With this many-layer model in mind, there are plenty of knobs to turn to let more and better Babble through.

The River of Babble

Imagine that your river of Babble at its source, the subconscious: a foaming, ugly-colored river littered with half-formed concepts, too wild to navigate, too dirty to drink from. A quarter mile across, the bellow of the rapids is deafening.

Downstream, you build a series of gates to tame the rushing rapids and perhaps extract something beautiful and pure.

The First Gate, conscious thought, is a huge dam a thousand feet high and holds almost all the incoming thoughts at bay. Behind it, an enormous lake forms, threatening to overflow at any moment. A thick layer of trash floats to the top of this lake, intermixed with a fair amount of the good stuff. The First Gate lets through anything that satisfies a bare minimum of syntactical and semantic constraints. Thoughts that make it past the First Gate are the first ones you become conscious of - that's why they call the output the Stream of Consciousness.

A mile down the Stream of Consciousness is the Second Gate, spoken word, the filter through which thoughts become sounds. This Gate keeps you from saying all the foolish or risqué thoughts tripping through your head. Past the Second Gate, your spoken words form only a pathetic trickle - a Babbling Brook.

By now there is hardly anything left to sift from. The Third Gate, written word, is no physical gate but a team of goldpanners, scattered down the length of the Babbling Brook to pan for jewels and nuggets of gold. Such rare beauties are the only Babble

that actually make it onto paper. You hoard these little trinkets in your personal diary or blog, hoping one day to accumulate enough to forge a beautiful necklace.

Past the Third Gate, more Gates lay unused because there simply isn't enough material to fuel them: a whole chain of manufactories passed down from the great writers of yore. Among them are the disembodied voices of Strunk and White:

> Omit needless words. Vigorous writing is concise. A sentence should contain no unnecessary words, a paragraph no unnecessary sentences, for the same reason that a drawing should have no unnecessary lines and a machine no unnecessary parts. This requires not that the writer make all his sentences short, or that he avoid all detail and treat his subjects only in outline, but that every word tell.

Jealously clutching the 500-word pearls you drop once a month on your blog, you dream of the day when the capital comes through and these Gates will be activated to produce your magnum opus, your great American novel. For now, you can't afford to omit a single precious word.

The Gates of Prune

In the model above, there are many problems with Prune independent of having low-quality Babble to begin with. The Gates are working at odds with each other. They are individually too strict. There are simply too many of them. Lots of expensive mental machinery is not working at full capacity, if at all: if you have four Gates but 99% of the goods don't make it through the first one, that novel-writing factory you've built is not paying rent.

Even worse, there's probably two or three layers of subtlety within each of the big Gates I sketched. What you might whisper on a dark night in total solitude is different from what you might utter to a confidante is different from what you might say to your thesis adviser.

If a balanced Babble and Prune game is supposed to involve one Artist against one Critic, then having an overactive Prune is like pitting a pitchfork-wielding mob of Critics against one Artist. The first three Critics tar-and-feather the Artist and the rest are just there for moral support.

The task of relaxing all of Prune at once is monumental. Instead, relax the Gates individually in order. Simultaneously, shorten the psychological distance between them.

Relaxing and Shortening

At the First Gate, conscious thought, noticing is the way to let through more subconscious Babble. Practice noticing thoughts and sensations (not just confusion) that you never pay attention to. Much of meditation is devoted to relaxing this first Prune filter. Much of art is devoted to the motto: *make the familiar strange*, where strange is better translated as *salient*.

Another exercise along similar lines is zooming in on anything, anything at all. Pick up and stare at the whorls and aphids running down that twig on your driveway. Take

apart that broken old Canon in the attic. Dissect your aversions toward attending Algebraic Geometry.

At the Second Gate, spoken word, the trick is getting comfortable with vocalizing more of your Stream of Consciousness. I mentioned before that my internal process is very verbal - on reflection I think that whole post is about the maturation of my Prune filter to allow more Babble through. Several features stand out.

One of these features is that I directly mouth or whisper any thoughts that appear in my Stream of Consciousness. Psychologically, this shortens the distance between the First Gate and the Second Gate: it becomes a question of how loud to speak rather than whether or not to speak at all. There's no reason not to be constantly mouthing the things you're thinking, at least when you're alone. Similarly, when lost in thought I make micro-gestures with my fingers to imitate the emphatic ones I would make to convey that point in conversation. These tricks exploit the fact that the psychological distance between 1% and 100% is much shorter than that between 0% and 100%.

Another feature of my internal process is that I always have a mental audience: a silent judgmental muse, the personification of the Critic. In HPMOR, Harry has a supersized version of this: a whole cast of colorful mental characters that carry out full-length conversations with each other. This kind of dissociation-into-subpersonalities exercise has a whole of great side effects, but the relevant one for us is that it again shortens the mental gap between the First and Second Gate by making thinking feel like conversation.

Onwards to the Third Gate: the written word. Thankfully, modern technology has already radically shortened the distance between the Second and Third Gates for us with the invention of the blog, a medium much more free-form and personal than the book. Your training as a writer has probably erected a tall Third Gate, and successful bloggers have pretty much circumvented it.

What distinguishes blogging from formal writing? One metric is the frequency with which the blogger breaks the Fourth Wall - that poor Wall which is only mentioned when it is broken. Having torn down the Fourth Wall, blogging reduces naturally to a heated and cogent form of conversation, filled with rhetorical questions and injunctions.

Hey, look here, I'm not saying there's no place whatsoever in writing for formality. But if you're going to build a wall and call it the Fourth Wall, build it after the Third Gate, you know?

Circumambulation

This is a linkpost for
https://radimentary.wordpress.com/2018/01/15/circumambulation/

This is Part 4 of the Babble and Prune sequence.

In the previous parts, I described the brain's thought-generation process as an adversarial learning system between Babble - which generates low-quality content - and Prune - which filters for high-quality content.

My primary motivation for understanding this system is to solve writer's block. On some days, I achieve a state of flow and write two thousand lucid words in two hours, which somehow coalesce a previously disconnected body of thoughts. On others, I give up after one or two navel-gazing abortions. These irregularities are a sign that my model of Babble and Prune is insufficiently predictive.

Babble and Prune are part of the picture, and a conscious effort of systematically relaxing the lower Gates of Prune has helped me produce more material. However, there's at least one important point which is entirely missing in this model: the choice of subject matter, the focus of one's attention towards which Babble and Prune approach.

Today I will fit Babble and Prune into another piece of the puzzle: Circumambulation - walking around the truth, spiraling towards it.

Truth is a process of successive approximation. More precisely: truth is a process of circumambulation. Deep truths are not mere point masses on a line to be approached with binary search or gradient descent. They are sprawling manifolds in high-dimensional space, and to map out such a Titan requires patience, false starts, and approach from many directions.

Circumambulation is the underlying directive of Babble, the process of circling around the holes in your understanding to pin down their shape and fill them in with substance.

1. Circumambulation

From Jung:

> I began to understand that the goal of psychic development is the self.
> There is no linear evolution; there is only a circumambulation of the self. Uniform development exists, at most, at the beginning; later, everything points toward the centre.
> This insight gave me stability, and gradually my inner peace returned.

This quote is accompanied by the lovely picture of one of the labyrinths found in many cathedrals. Classically, cathedrals are built in the shape of a cross, e.g. Notre Dame:

At the center of the cross (X marks the spot) is the labyrinth. Walking the labyrinth is a ritual reflecting on the nonlinearity with which we circle the truth.

Our imagery for reality is deeply tied to our spatial reasoning. Metaphors like map and territory, thingspace, and pendulums make instinctive sense by tapping into this connection. New knowledge is produced by honing in on holes in the map, and then filling them in by surveying the territory.

The map can be a very twisted and high-dimensional thing full of holes and jumps, and Babble is a biased random walk like simulated annealing on this space.

Babble is the process of randomly walking around the map, honing in on beacons, holes, and discontinuities. These are (at least) three different ways in which something can be missing from the map. Properly done, Babble can detect all of them.

i. Beacons

A beacon attracts attention. Like the summit of a mountain or the base of a valley, it has the feel of a local optimum. Circumambulation in known regions of the map spiral inevitably into beacons - recurring thoughts and preoccupations.

Beacons signal there is more to the story - the single marker on the map that you knew lazily as "procrastination" is actually a whole category of distinct mental phenomena. Once a beacon is identified, it needs to be blown up because there is too much information to be stored at a point. Imagine surgically expanding the beacon into a whole bubble of space.

Newcomb's problem is still a beacon for me. Every time I wander into its general region, I add a bit of detail to my understanding. I gain a little more respect for credible precommitment and iterated games and a little more certainty of one-boxing in superficially Newcomb-like problems in real life. Nevertheless, a beacon of residual unease remains.

ii. Holes

Holes in the map are gaps in your knowledge and models. You can infer the existence of simple and low-dimensional holes locally - they are usually marked by weak beacons. Knowing plenty about how baby humans develop, you infer that there is just as much detail to learn about crocodiles. Babble can sometimes lead you to fill in these simple holes out of curiosity, and find out that crocodile gender is determined by the temperature of the egg. Nevertheless, you don't have time to fill in all the little gaps.

Big gaps in your knowledge may be so big you don't even know they exist. If you walk in a straight line while staring at your feet, you'll miss most of the serious holes in the map. That's the value of circumambulation: detecting holes more globally when local detection is difficult or impossible.

The map may be a hollow sphere (or worse, a flat torus) that needs to be filled in. Walking around myopically, you think you live in Flatland. But if you Circumambulate mindfully, you might notice your Babble looks like this:

Detecting that three right angles don't make a triangle, you notice that your map has positive curvature, and there might be important insight at the center of this spherical

hole you've been skirting around. Insight that you couldn't have found by walking in straight lines.

iii. Discontinuities

The confusion you notice as you Babble can signal holes in your knowledge, but it can also mean there are inaccuracies in what is already there. Figuring out what to do with a note of confusion is a delicate science. That slight pain in your chest - is it heartburn or lung cancer?

You make a circle around the concept, only to end up not where you started. Maybe it's a sign of a hole, or lack of precision. You think the map is a circle, but actually it's a covering space thereof, i.e. a much bigger circle. Your models are not sufficiently precise to distinguish subtle differences between points in the same fiber. It might also be the sign of a discontinuity, an internal contradiction in your map. Imagine the dismay of a cartographer who tries to build an atlas based on Columbus' reports of reaching India, unaware of the entire New World.

Discontinuities can be global just like holes, and require global solutions, and to rebuild models from scratch for a single discontinuity is usually too expensive. Identifying and eliminating these systematic biases - or at least correcting for them - is a central rationalist project.

2. Circumambulation as Battle Royale

Previously, I compared Babble to sampling from Google's PageRank algorithm: taking random walks on the Babble graph with random restarts. The randomness is biased by a weak, local heuristic.

But however useful PageRank is, it's only once in a blue moon that someone wants a global list of all extant web pages sorted by popularity. Google needs a search term and filters to be useful. Similarly, every Babble walk has to start at a basepoint, and it's less random than I suggested.

Narrowing down to a central truth is like a game of Battle Royale. Initially, many different hypotheses are spawned all over the map and fan out in pseudo-random walks. However, as Babble generates hypotheses and some of them are Pruned, we begin to triangulate the coordinates of the center. The safe area shrinks.

Babble is produced starting from a basepoint and restricted to a general region. As a clearer picture of the center is developed via Babble and Prune, the basepoint shifts closer and closer to it, while the region of interest shrinks like the safe area. These constrained random walks are forced to clash, and stronger hypotheses eliminate weaker ones or coalesce until at the very end, only one winner remains.

3. Focusing is Circumambulation

This recent post on focusing fits very cleanly into Babble and Prune. I recommend reading the whole post, but here are some relevant points:

> The "big idea" of Focusing (according to me) is that parts of your subconscious System 1 are storing up massive amounts of accurate, useful information that your conscious System 2 isn't really able to access. There are things that you're aware of "on some level," data that you perceived but didn't consciously process (see blindsight as both concrete example and metaphor), competing goalsets that you've never explicitly articulated, and so on and so forth.
> [...]
> (There I was doing a super-fast scan over a whole bunch of possible words and phrases and explanations, all the while paying very close attention to my truth-detection module. I knew where the answer *would* be, and I knew its general shape, but I had to keep looking until I found something *juuuuust* right. It's like when you mentally stutter past five or six different comebacks to throw at your sibling until you find the one that's cutting, true, *and* okay-to-say-even-though-your-parents-are-listening.)
> You get the idea. As the process continues, the picture grows more and more accurate, and *evokes* more and more of the underlying what's-really-going-on. I can feel a sort of click, or a release of pressure, or a deep rightness, once I say the thing that *really* completes the picture.

In our language, massive amounts of accurate, useful information are stored by System 1 behind the First Gate of Prune, and focusing is a particular exercise towards

noticing previously subconscious information (beacons, holes, and discontinuities) and letting useful Babble through. Prune is a necessary part of this process, killing weak hypotheses and directing Babble in the proper area.

I have described a general, not entirely novel, approach to tending your mental garden, circling around the map via Babble in search of precision, completeness, and internal consistency. The upshot of this post is that Circumambulation is the right geometric imagery for this process, that geometric metaphors highlight the possibility of serious and nonlocal flaws in the map, and therefore that spiralling around the truth at length is rarely a waste of time.

Write

This is a linkpost for https://radimentary.wordpress.com/2018/01/28/write/

This post is Part 5 of the sequence on Babble. After writing Hammers and Nails, I figured out that my favorite Hammer is writing.

Write about everything. Write it immediately. Edit afterwards.

In this post, I tell two stories about the magic of putting words on paper. Make your own conclusions about the brain-enhancing effects of writing.

Then, I share my recursive blogging process which avoids the ickiness of planning and gets directly to the thick of things. The recursion makes it easy to expand sections without breaking into structureless rambles. My process is mostly inspired by Nonfiction Writing Advice.

Finally, I explain why disclaimers suck.

Meaning Injection

In middle school, I carried a single enormous binder to every class, filled with colored tabs and lined paper. In that binder, I took notes religiously, trying various techniques of organization and highlighting.

I had no good notion of why I took notes. After all, I never looked at the notes I took - textbooks were always easier on the eyes than my chicken scrawl. I did notice, however, that I remembered things better when I wrote them down, so for a time my plan was to simply take the notes and forget about them in my binder.

I eventually noticed what was actually working, and discarded the superstition. I left my binder to gather dust and started bringing lined paper to class to take the notes. When the bell tolled, I dropped the notes in the recycling before leaving the classroom. I wrote down only the key words and spent the rest of class staring at those scattered words, injecting them with the meaning. Imagine lonely crimson ink-drops falling into a glass of water, bursting into delicious, velvety tendrils at contact. Plop, Plop. That's how meaning injection feels.

Several iterations later, my study strategy is perfectly streamlined.

I no longer take notes in class. The night before each test, I skim the textbook with a blank sheet of paper and write down all the important terms - without definitions; the sheet of paper serves as a kind of corporeal mind palace. I read through the list once, injecting the words with meaning. Plop, Plop. And that would be enough.

Research Notes

Fast forward to 2013 and transport yourself to my first summer research program. Every Monday, we give a brief board about that week's progress. I mull ideas on paper over the week before TeXing them up Sunday night.

A curious thing happened - all my progress happened on Monday and Tuesday. I spent the rest of the week meandering around the same ideas, checking special cases and writing up fragments of arguments. On Sunday night I write everything down, and the ideas crystallize on paper. They lose their grip on me, and I move on to new pastures.

These days, I reproduce the effect simply by writing up any partial results as quickly as possible. Oftentimes, I sit down with an uncertain inspiration and immediately try to TeXing it. Usually, it fails catastrophically but I gain a piece of insight. If successful, the argument pops out on paper fully-formed by narrative force. My last paper - itself only twenty pages - is a Frankenstein-esque construct built out of thirty rambling TeX documents.

Structure is Process

Very readable and natural blog posts can be written recursively with minimal pre-processing. My process is mainly designed to streamline brainstorming and outlining while avoiding two failure modes.

The first failure mode is simply writing too little. Thus, I emphasize rapidity over content and quantity over quality.

The second failure mode is writing long, detailed pieces with little structure, and finding myself disagreeing from the beginning of the argument by the time I reach the end. You know the loudness bar over the microphone icon in Skype settings? The length of my blog posts used to fly up and down just like that capricious little bar. I'm incapable of finishing two paragraphs without changing my mind, and my process is designed to recursively deconstruct into semi-independent pieces that I can close and set aside.

1. Headers are Brainstorming

Brainstorm by breaking your argument down into a handful (2-5) of the core ideas. Separate individual arguments and anecdotes. Write down titles for each.

For this post, I started with four sections: "Study Skills," "Research Notes," "AI Risk and Affordance Widths," and "Structure is Process." Later on, section 3 was removed - I couldn't find an angle to tie AI Risk with the magic of poetry. Also, "affordance width" is a wonderful idea, but the concept handle screeches like cheap chalk. Would someone please syllable it?

2. Intros are Outlines

The introduction to a post defines the topic and sketches the arc of each section. It teases the reader with sneak peaks. Your post is cobbled together from unrelated shit but the reader won't notice if you tie it together with a metaphor or something.

In the introduction, it can be useful to gather links to background reading and place the post in a larger context. The reader finds this thoughtful and welcoming. Actually, it's a convenient trick to batch-process all the link hunting.

3. Recurse

Recursively apply this process down to the level of sentences. Each section longer than a single complete thought should be broken into subsections and outlined. Each paragraph longer than five sentences should be broken into smaller paragraphs.

At the lowest level, it may be wise to treat topic sentences as micro-headers, to be written first before filling out the paragraphs with content. It may also be wise to reflect on the integration of the (sub-)sections into the whole before filling them out.

On Disclaimers

If you think your words are but pale wavering shadows of the real content in your heart, you are not alone. You will feel the need to precede your posts with all manner of content warnings, epistemic statuses, and protests of humility - you will feel no right to cry black and white when all the world is shades of gray.

There are decent and genuine reasons for disclaimers, but I'll let you produce those. I propose that we use epistemic statuses and disclaimers sparingly.

What happens to the reader when every post starts, "epistemic status: mostly true with a chance of rain?" As the reader's eyes hit "epistemic status," notice how they roll to the side and glaze over ever so slightly.

Ugh.

She knows it's 10% about epistemic and 90% about status. Magic brain juice [citation needed] takes over and associates that instant "ugh" with opening your posts.

Instead of conditioning readers to hate us, I propose we return to a saner time, where the fact that your words are but a pale wavering shadow of the grand, mysterious truth in your heart is *the default assumption* about human communication. Where truth is a dance of successive approximations yet no step in that dance requires adult supervision. Where quibbles over certitude are banished to the comments section where they belong.