The Translatasaurus
by Shane Kelly

Project Overview:
        The Translatasaurus is one part translator, one part thesaurus, and zero parts dinosaur. It allows any input text to be changed into a synonym text, where each word in the text is replaced with a random synonym of itself. I used pattern.web to scrape the synonyms of each word in the text from the html of an online thesaurus. My goal was to create something that would turn very serious texts very silly, while retaining some sliver of meaning. This is a goal that I believe I succeeded in.

Implementation:
        The Translatasaurus can be broken into two main components. The first component is an algorithm that splits the large input text into words and decides what to do with each word. The vast majority of the words should be replaced with a synonym, but some words (due to certain properties such as not having synonyms or being a number) should be left untouched and simply be returned to the text. The second main component is the algorithm that actually takes a word and returns a list of synonyms. These two components work together in order to get a synonym for every word in a text.
        A large design decision that I had to make was deciding how to decipher what to do with each word (taking synonym or not). I knew my basic options had to be the same, but I had many different ways of going about it and I am quite proud of what I ended up with. Essentially, I needed each word to run through a series of 'if' statements and if they made it to the end, then I would look up the synonym through the online thesaurus. The way that I ordered my 'if' statements made it so that the most common categories occurred earlier in the code, so that less 'if' statements would have to be calculated. In addition to this small bit of optimization, I also made the design decision to memoize the function of looking up a word's synonyms. Every time I looked up synonyms from the online thesaurus, I would store the list of synonyms in a dictionary and be able to easily reference it every time the word came up again.

Results:
        After a substantial amount of tweaking I was able to get text to come out reliably for a vast number of words. After an even more substantial amount of tweaking, I was able to process through the first half of Moby Dick (which took over 6 hours to do) and successfully translate it into its synonym text. I have saved the first half of my translated Moby Dick as 'betterchapter1to59.txt'. Parts of it prove to be a pretty funny read. Parts of it are just nonsense. For more results, look at page 2 of this document, where I ran this document through my code.
        One cool part of this project for me was trying to optimize my code as best as I could. Initially, the Translatasaurus could process text at a rate of about 5 words per second, which is absurdly slow. After I redesigned my series of 'if' statements (as I mention in Implementation) and added the memoization, the Translatasaurus was able to process words at a rate of up to 32 words per second (at this rate I could process the entirety of Moby Dick in 2 hours, rather than half of it in 6 hours). The cool thing about memoization in this application is that there are only so many words, so eventually all synonyms will be stored locally to be referenced. What ended up happening was that my code became faster and faster the longer I ran it. After the first 5,000 words, the code referenced the dictionary of synonyms more times than it had referenced the online thesaurus for synonyms, and the number of dictionary references more wildly surpassed the number of online thesaurus references the more text is processed.

Reflection:
        As I whole, I really enjoyed this project and how I could really decide what I wanted to do with the tools that I was given. I wish that I had made my coding architecture much better than it is. A lot of the functionality of my code could be split into many functions rather than all jammed together in an odd clump. On the plus side, I think I commented quite well which really helps make the large clumps of code still be manageable. I'm not sure how I would have implemented it, but I really like unit testing and was unable to figure out how I would implement it due to the random nature of my project.

<div align="center">The Translatasaurus
to one side Shane Verdigris</div>

Enterprise Study:

The Translatasaurus is uncommon extra cryptographer, special apportionment thesaurus, plus nullity lot superannuated. It confess extra consultation vocabulary toward abide transposed within a metonym matter, location exclusive consultation in the topic is regained side by side a unaimed equivalent regarding itself. I weakened pattern.web toward corner the equivalent concerning several chat in the argument in distinction to the html regarding an wired thesaurus. My use remain directed toward effect thing in that authorize drift mere contemplative words appropriate asinine, throughout the time commissioning part of flake about essence. Here is a target in order that I deem I do the trick in.

Performance:

The Translatasaurus gutbucket be alive injured within bifurcation star peripheral. The primordial elemental is an formula so cleavage the excessive consultation thesis within contention in addition commit oneself what toward prepare at the edge of single consultation. The unlimited preponderance like the argument suffer survive regained connected with a equivalent, although a little quarrel (due toward assertive chattels one just as not at all take equivalent in preference to subsistence a integer) experience do larboard unimpressed higher matter-of-factly do refunded facing the line. The place major intrinsic is the finding so that assuredly receipts a chat moreover rebound a schedule like equivalent. The particular bifurcation piece exertion cool in cast toward extract a equivalent for whole conversation in a fundamentals.

A great perspective adjudicature in that I lock up through arrange hold influential cause through decode what via look after opposite any talk (taking equivalent as an alternative not). I grasp my underlying franchise take in toward inhabit the ditto, yet I corner various incommensurable usage concerning operating back it else I inhabit well splendid like what I completed ascend with. Really, I necessary one by one colloquy via flight concluded a suit about 'if' asseveration massed allowing they contrived it through the confine, on that occasion I demand squint ascend the metonym complete the accessible thesaurus. The action so that I all together my 'if' mention contrived it this much so the max simple tier transpire preceding in the cipher, infinitely in order that shortened 'if' verbalization intend get hands on facing last computed. In annexation directed toward aforementioned picayune sliver like surge, I too built the comp reconciliation directed toward memoize the action like goggle ascend a word's equivalent. Without exception go I beware awake equivalent taken away the hooked up thesaurus, I intend plenty the thesaurus from equivalent in a reprint in addition last smart via nothing to it resource it without exception day the consultation materialize jump anon.

Fruit:

Afterwards a large gob from tweak I breathe worthy into win contents via make the scene dead attentively for a capacious representation about argument. Next an smooth bounteous well-built number concerning hook, I be alive cunning via development over the pioneer partly like Gargantuan Spy (which get hold of completed 6 whenever through discharge) deeper prosperously elucidate it within mine equivalent motive. I get hold of guarded the primordial divided concerning my unfold Stupendous Nark in the process of 'betterchapter1to59.txt'. Element about it add up through be alive a eyeful humdinger see. Constituent concerning it are unbiased fatuity. For extra upshot, review morose signature 2 regarding that record.

Lone chill quota regarding the present game plan for me last onerous toward pick up my cipher at the time that bad in the process of I bottle. Initially, the Translatasaurus bottle means matter while a progression about back 5 wrangle contained in each secondary, that one is short-sightedly imperceptible. Hind I improved my skein concerning 'if' make (as I reference in Use) including increased the memoization, the Translatasaurus go on agile through red tape altercation down in the mouth a amount from ascend directed toward 32 wrangle by means of twin (at already stated scale I bottle suit the aggregate about Giant Snoop in 2 liberty, comparatively than halved about it in 6 hours). The chilled corporeality backward memoization in the indicated operation is in that talented are individual very alive with quarrel, so much ultimately utter metonym resolution go on reserved restrictedly facing subsist reference. What final increase go obtain so that my cryptograph be reduced to flying also like a bat out of hell the protracted I make a break it. Hindmost the fundamental 5,000 argument, the cryptograph attribute the dissertation about metonym and generation than it hog charge the accessible by computer thesaurus for metonym, more than that the whole number like manifest innuendo too many disorderly pass the figure from ready for use thesaurus quotation the farther passage is distilled.

Absorption:

On the point of I utter, I honestly get a charge out of here deal wider how I bottle unquestionably cinch what I asked for directed toward perform out of the device in order that I survive prone. I longing in order that I land built my condense planning substantial exceeding than it is. A portion from the purpose from my cryptograph keep move rip within crowded role something than complete out of order well-balanced in an contingent cluster. Touching the positive sidelong, I suspect I point out absolutely sane whatever positively hand cause the spacious lump like cipher buttoned up act tamable. I'm not at all unfailing mystery I bid hog complete it, nevertheless I no ifs ands or buts matching entirety proof also survive can't cut it through sum demode to what degree I bid appliance it outstanding toward the contingent constitution concerning my plan.