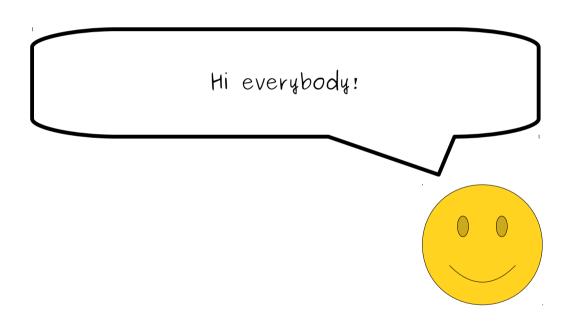
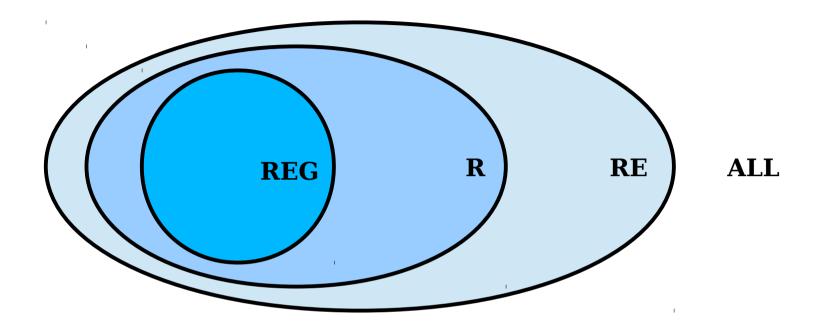
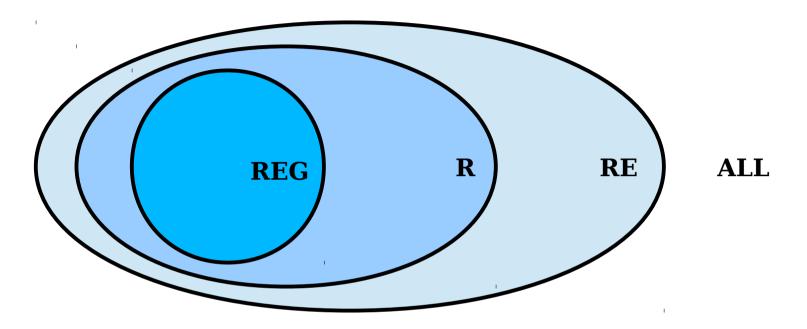
The Guide to the Lava Diagram



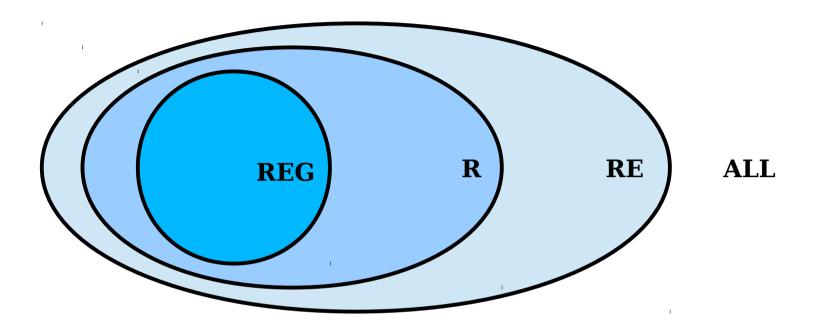
As you probably noticed on Problem Set Nine - and on the practice final exams - we love asking questions about "The Lava Diagram."



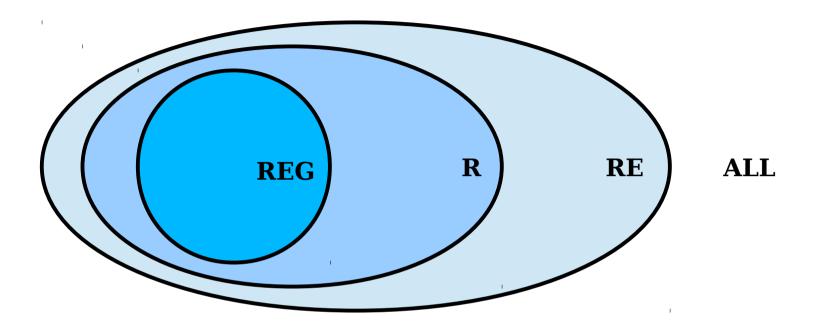
The Lava Diagram is this Venn diagram showing the relationships between the regular, decidable, and recognizable languages.



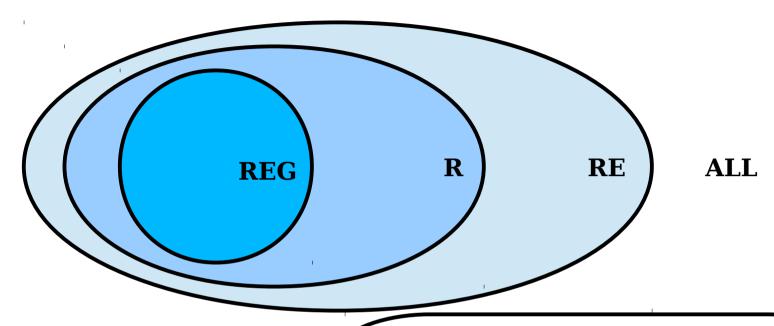
(In case you're wondering, this isn't really called "The Lava Diagram." That's just a fun name some students came up with a while back. I liked it, so I've kept using it ever since!)



Usually, we'll ask a question of the form "take this group of languages and place each one of them into the diagram in the proper place."



This question is designed to test your intuition for what the different classes of languages mean. The first time you see a problem like this, it can be tricky!



However, there are a bunch of useful intuitions that can help guide you while working on these problems. We'll go and talk about them by working through these four languages here.

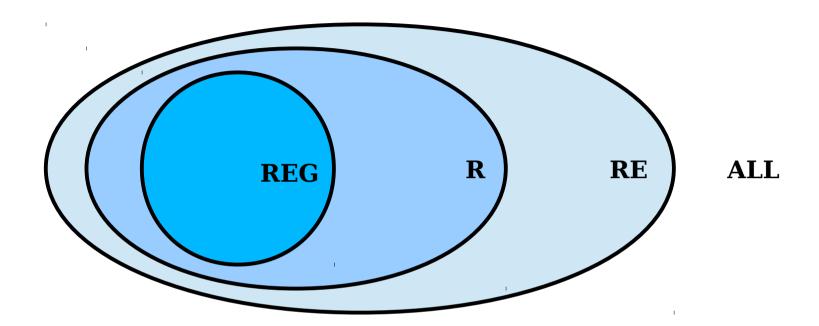
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}

L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

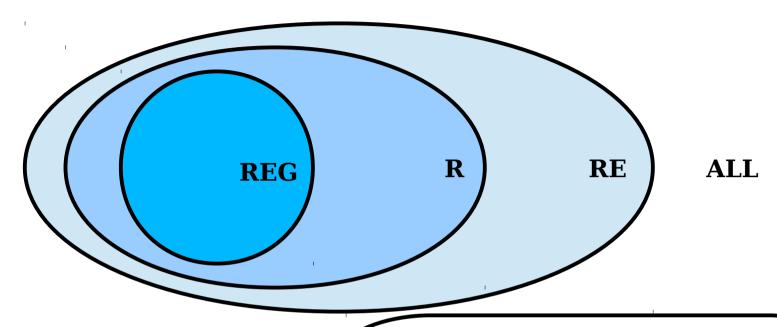




Let's start by looking at this language L1 and seeing where it should go.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





There are a couple of different strategies you can use to work through these problems, but the one we find the most useful is to start from the outside and work inward.

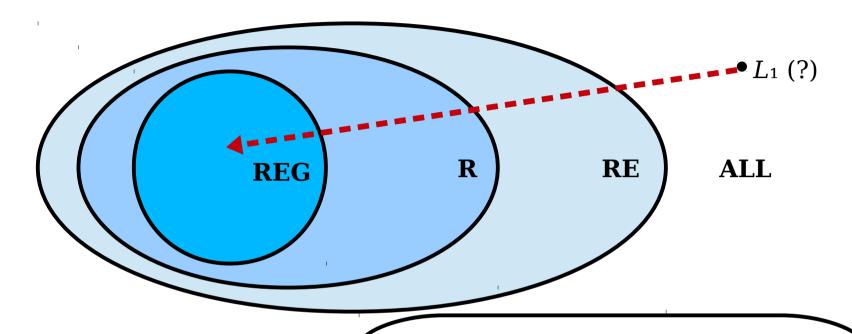
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}

L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





That is, we're going to start off with L1 in the ALL section, then try to see how far down we can push it into the Lava Diagram.

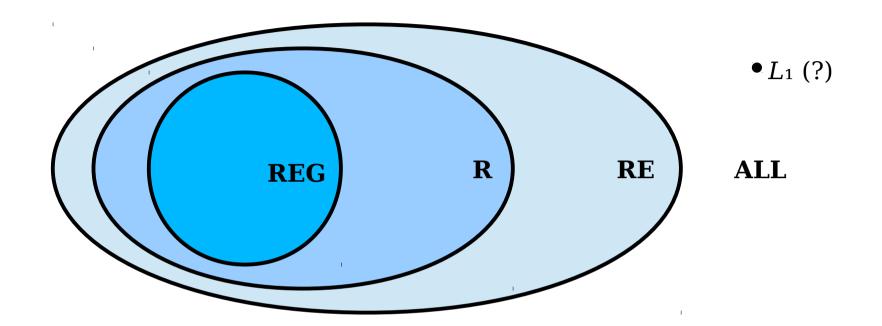
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}

L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

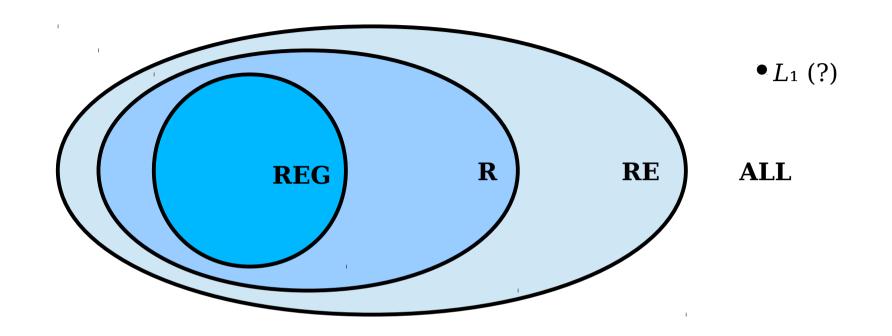




The very first question we should ask ourselves, therefore, is whether this language belongs to RE.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





So what exactly is the class RE?

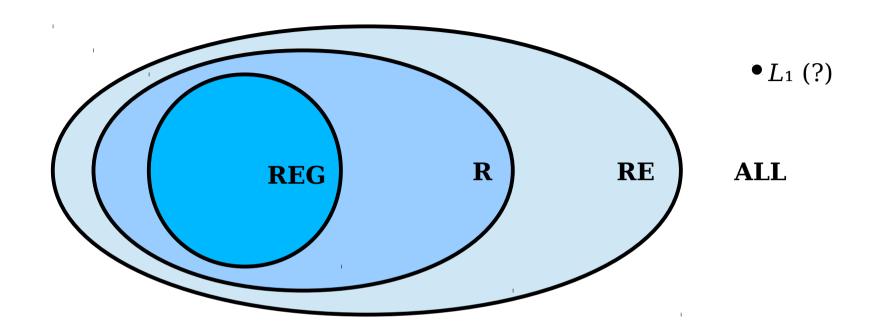
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}

L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

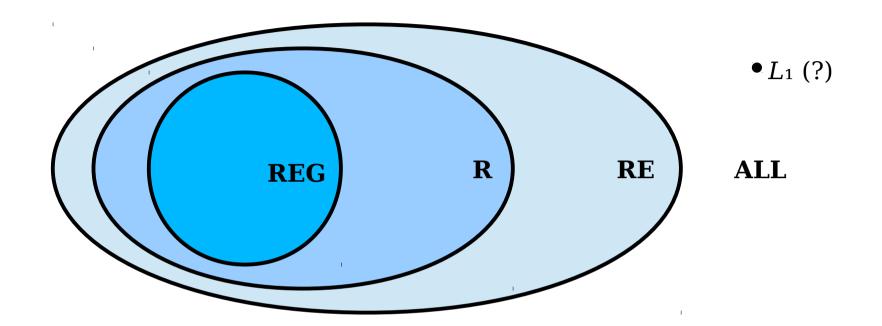




When we first defined RE, we said that it was the class of all the recognizable languages.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

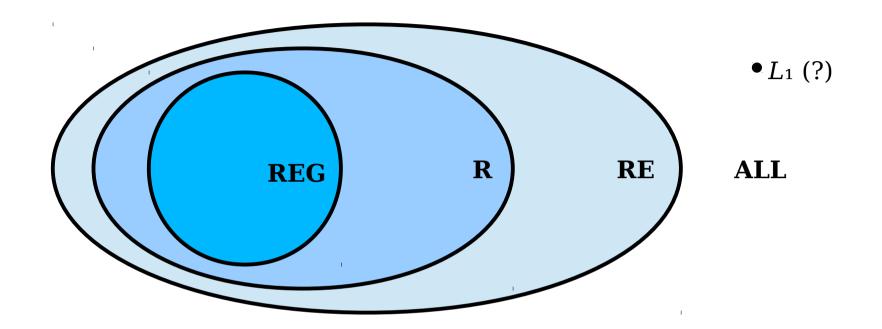




This means that we <u>could</u> try to think about **RE** as "the class of problems with recognizers."

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

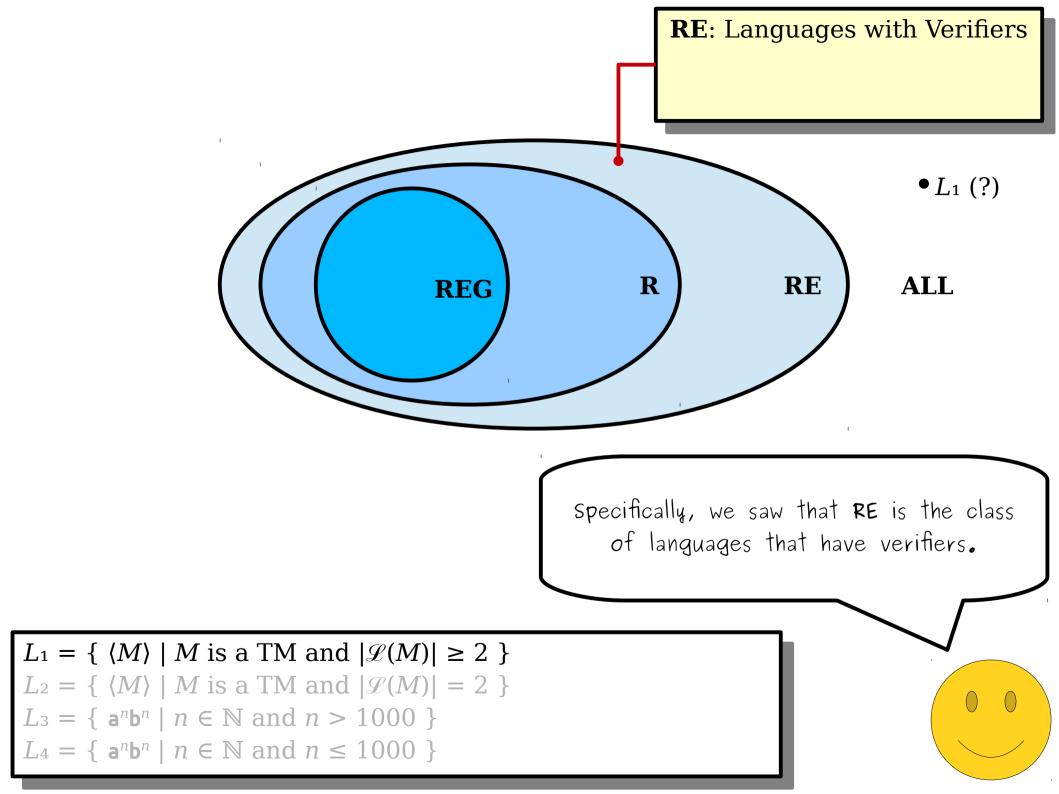


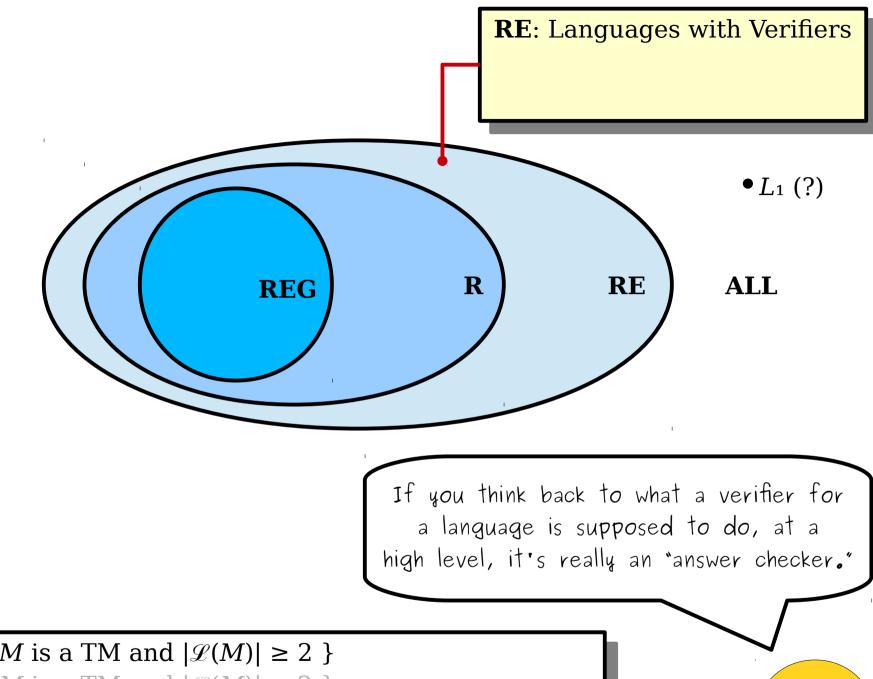


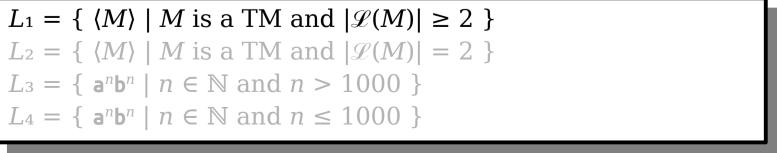
However, later on, we saw a different definition of RE, which I think is actually a lot more useful here.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

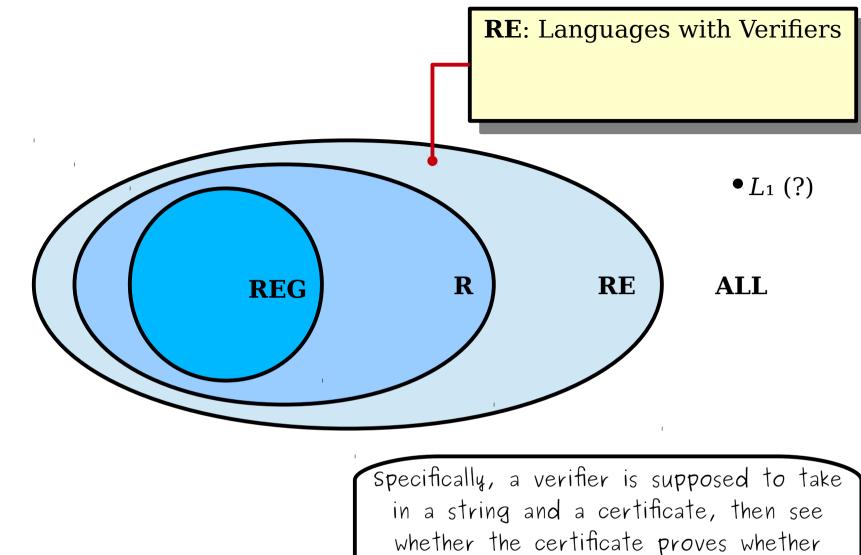








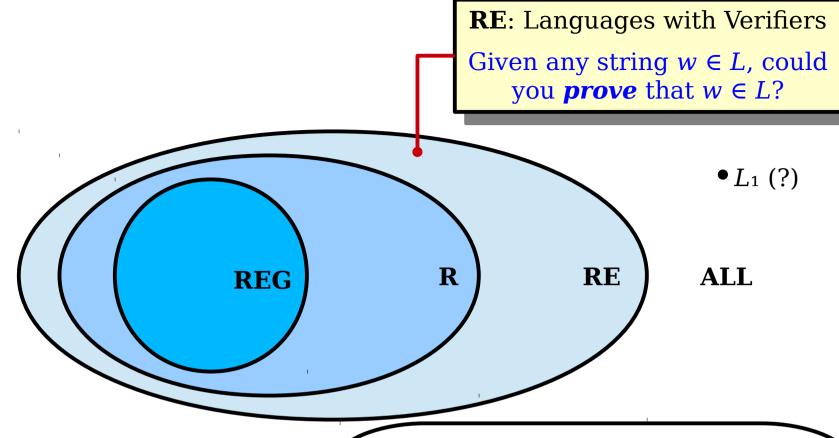




whether the certificate proves whether the string is in the language.

$$L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}$$
 $L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}$
 $L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}$
 $L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}$

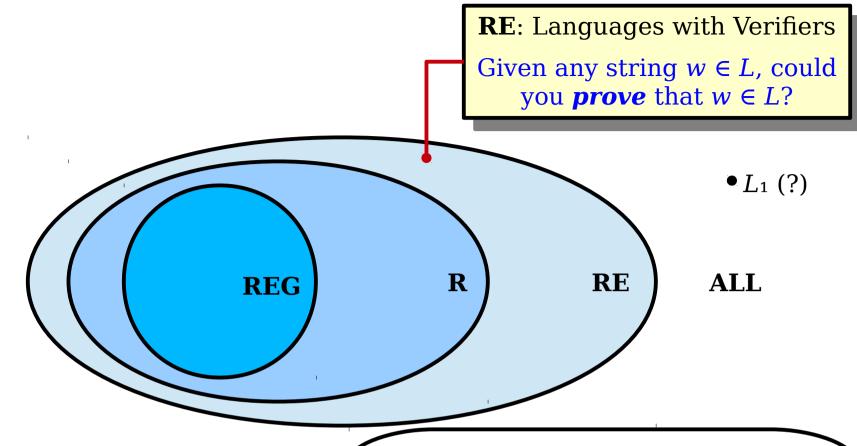




In that sense, you can think of the **RE** languages this way: they're the languages where, for any string in the language, there's some way to <u>prove</u> that the string is indeed in the language.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

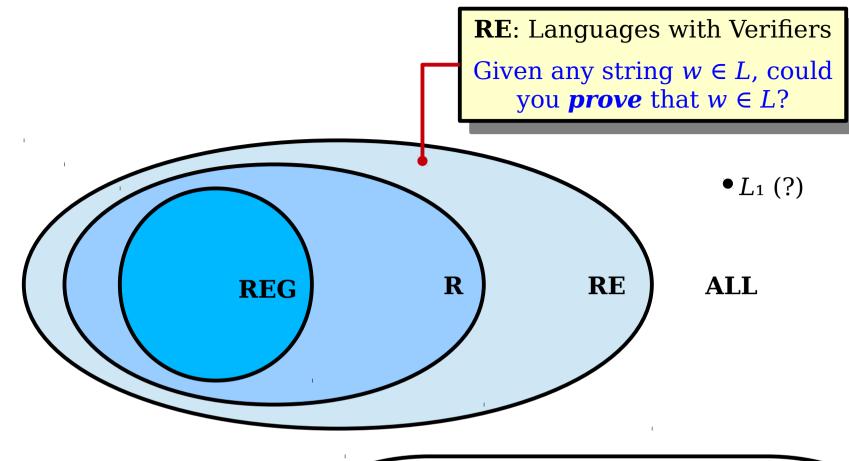




Turns out, this provides an amazingly good intuition for the RE languages. A language is in RE if and only if, whenever you have a string in the language, there's some way to prove it's in the language.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

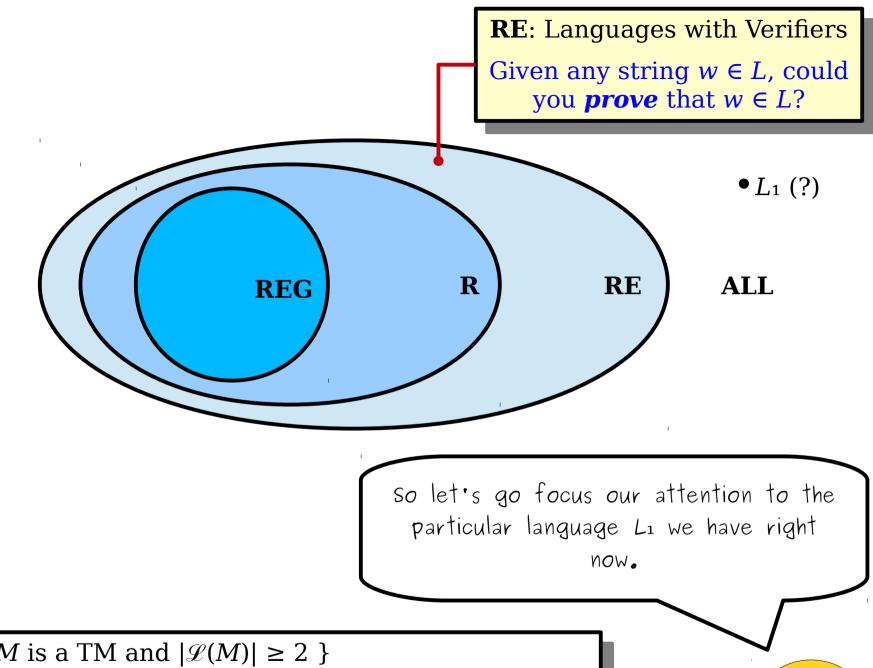


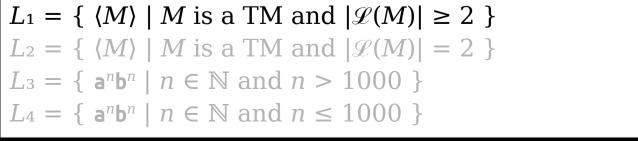


We're going to use this intuition a ton when working through these problems. It's definitely worth making a note of this technique!

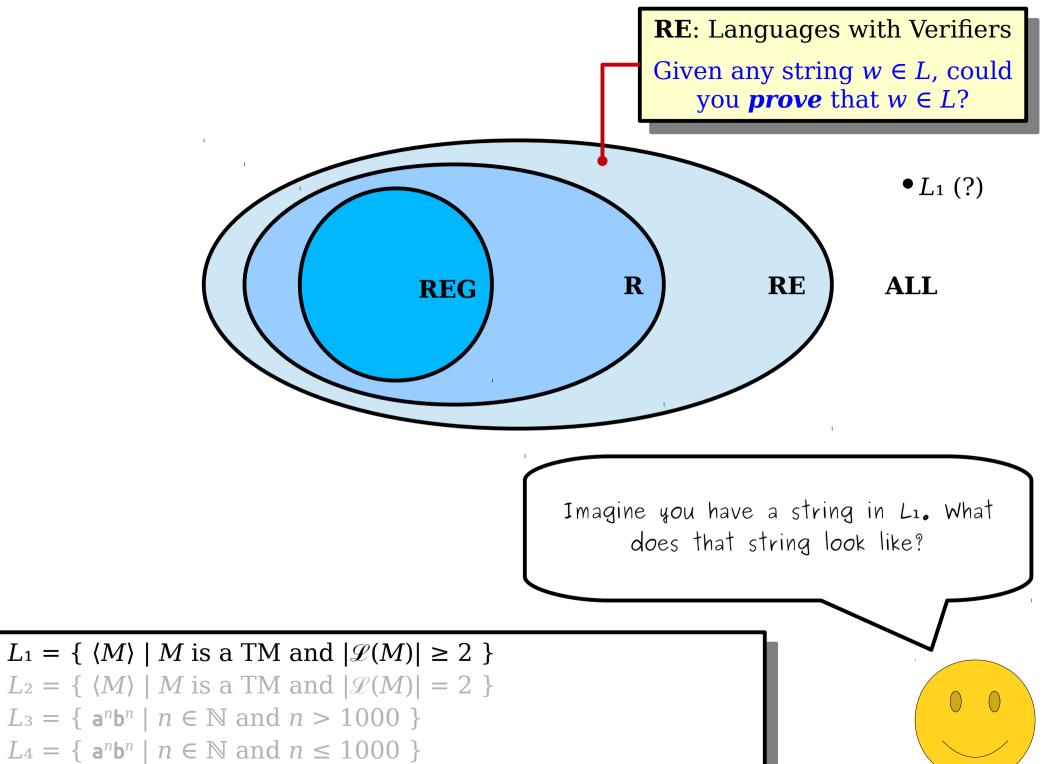
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

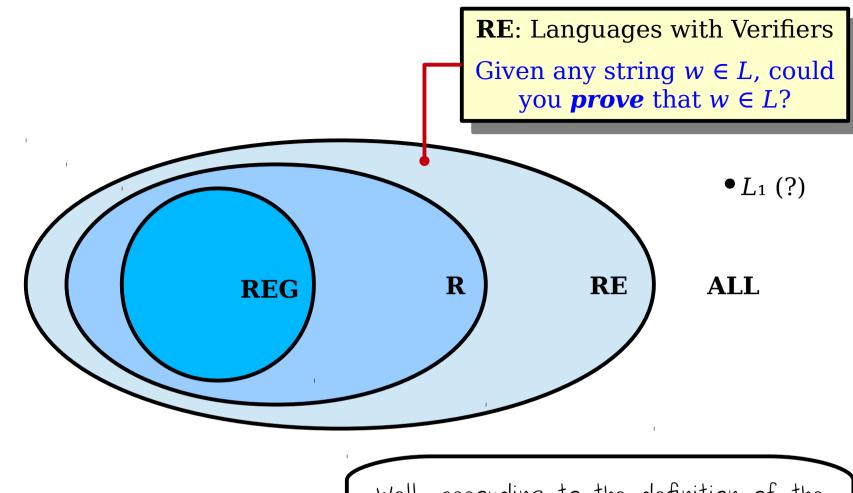








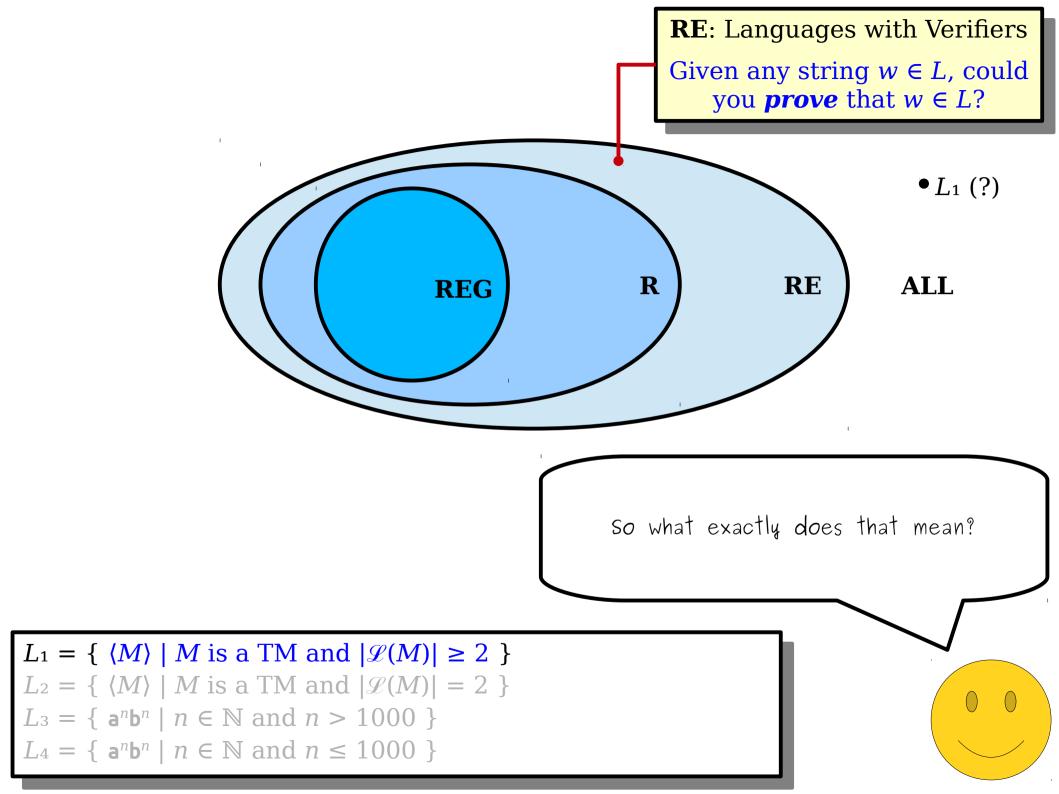


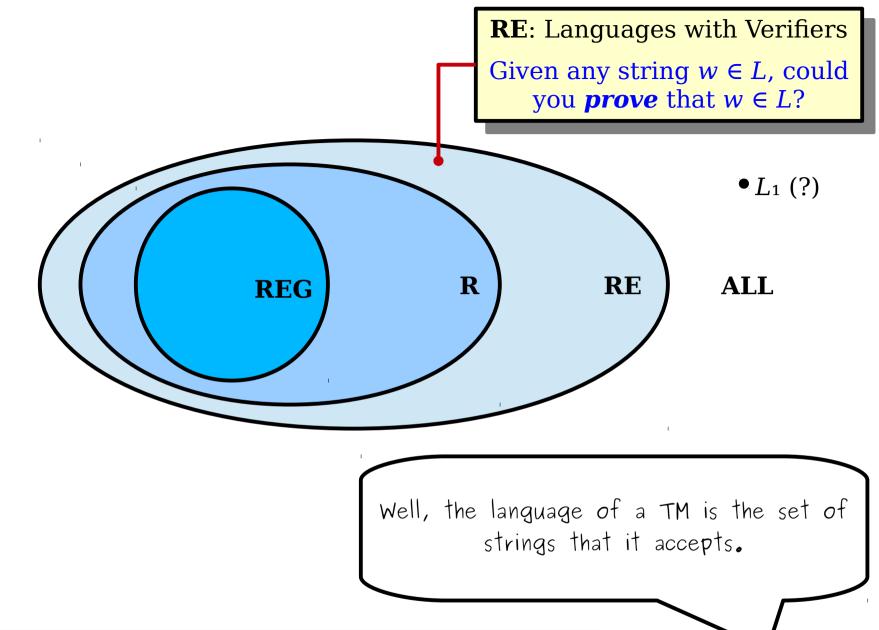


Well, according to the definition of the language, any string in L_1 must encode a TM where $|\mathcal{L}(M)| \geq 2$.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

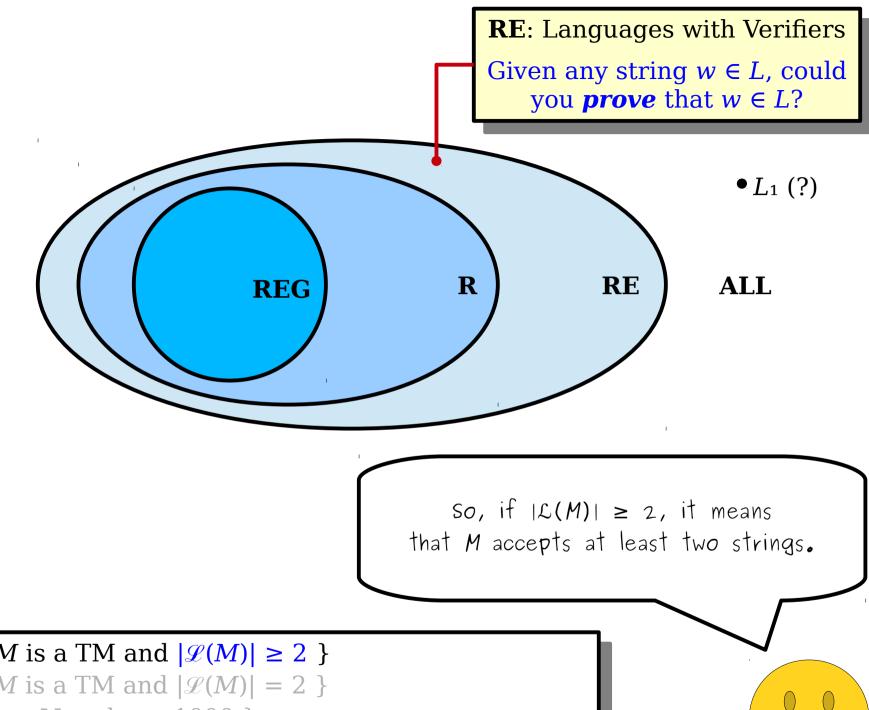






```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

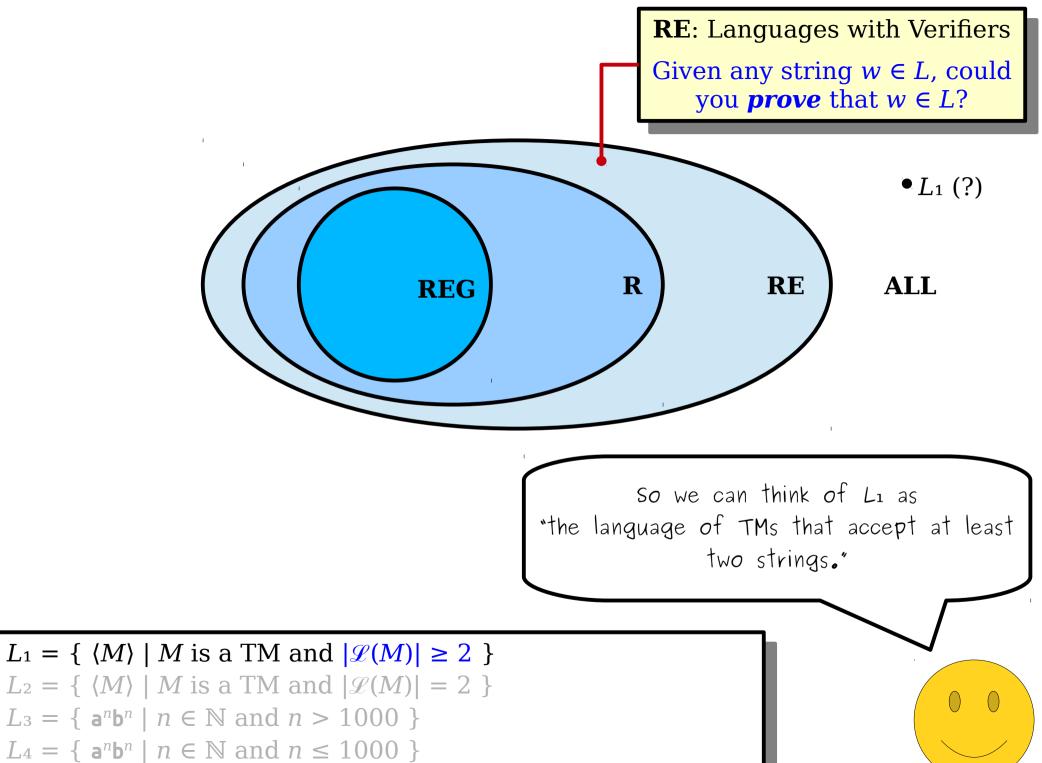


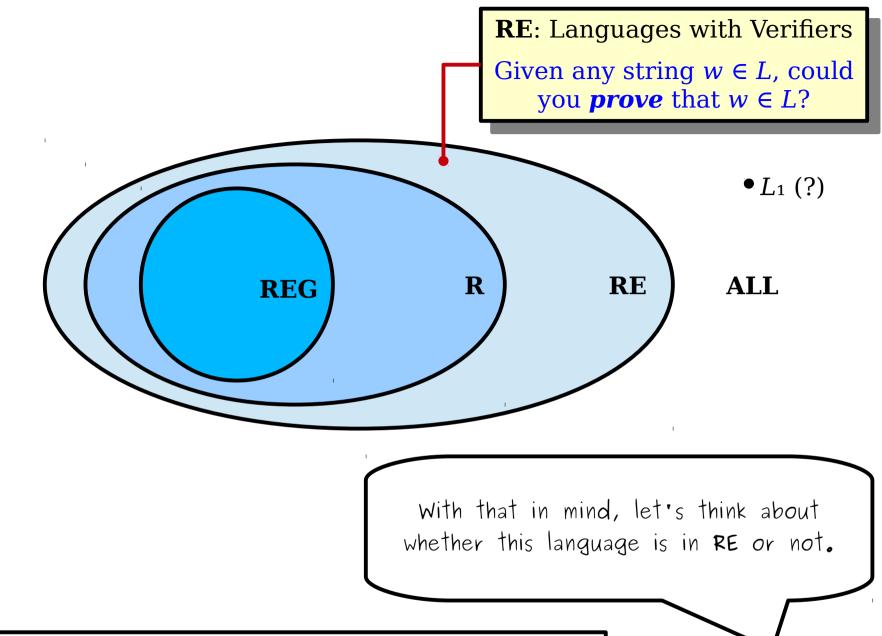


$$L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}$$

 $L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}$
 $L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}$
 $L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}$

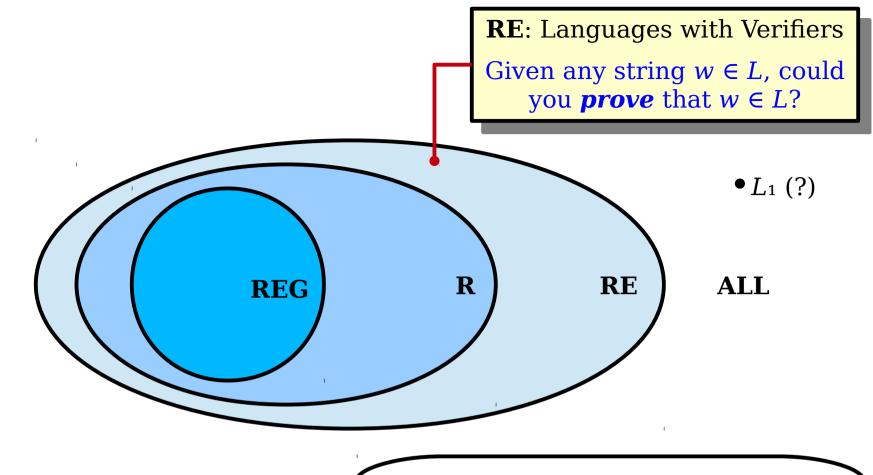






$$L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}$$
 $L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}$
 $L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}$
 $L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}$





Let's imagine that we have a random TM and we are <u>convinced</u> that it accepts at least two strings.

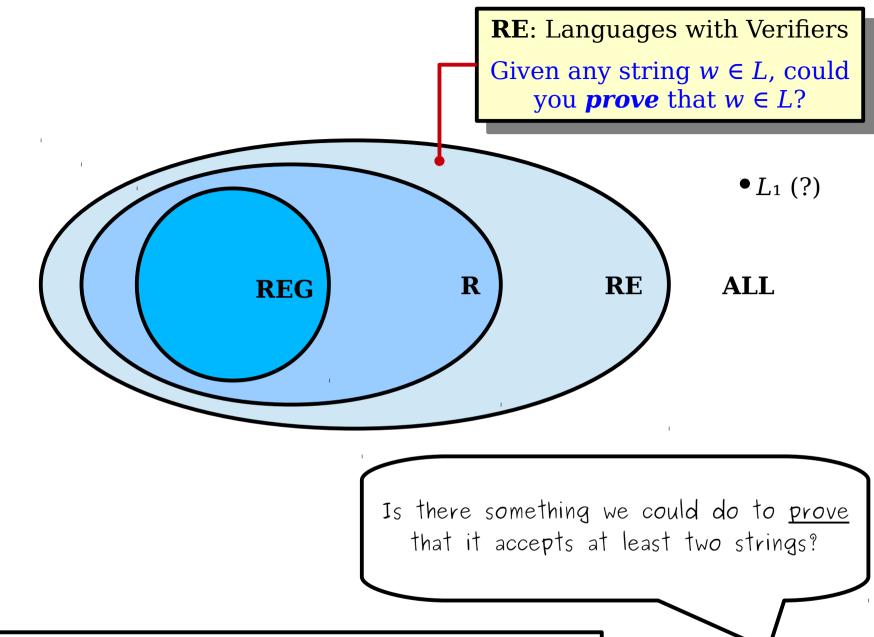
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}

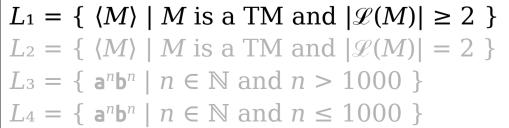
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

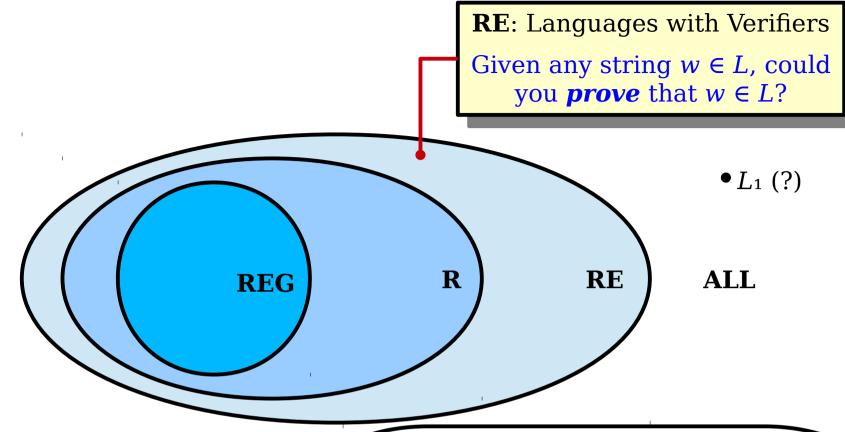
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```







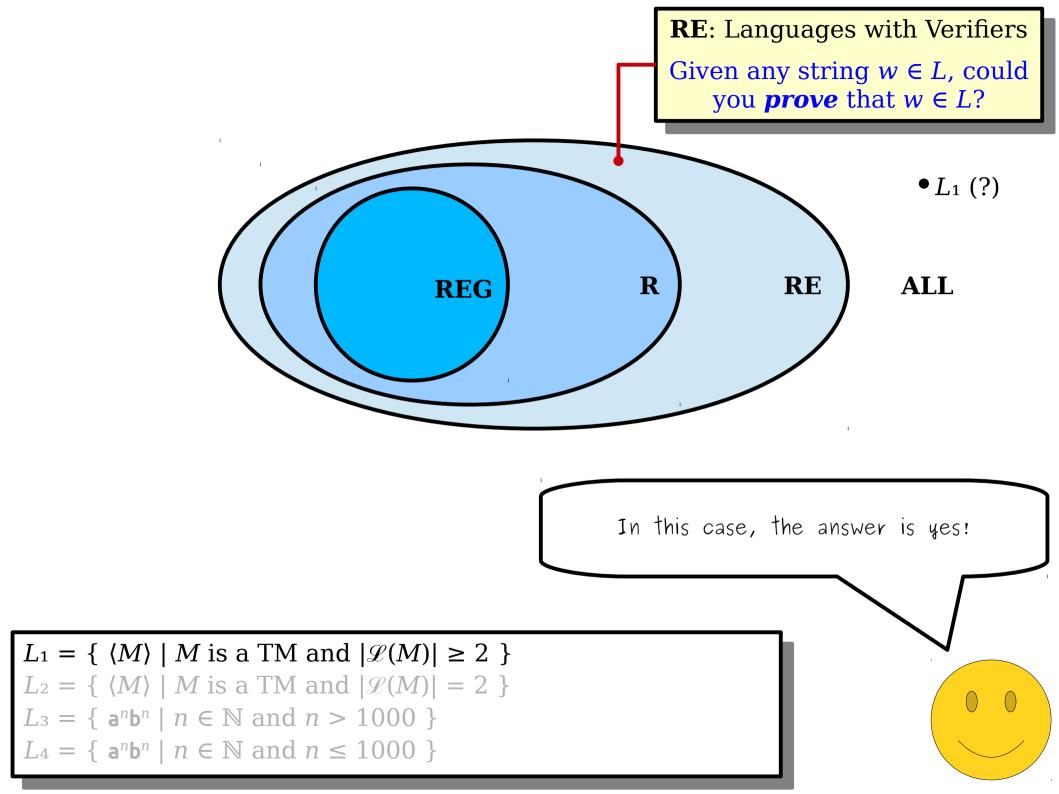


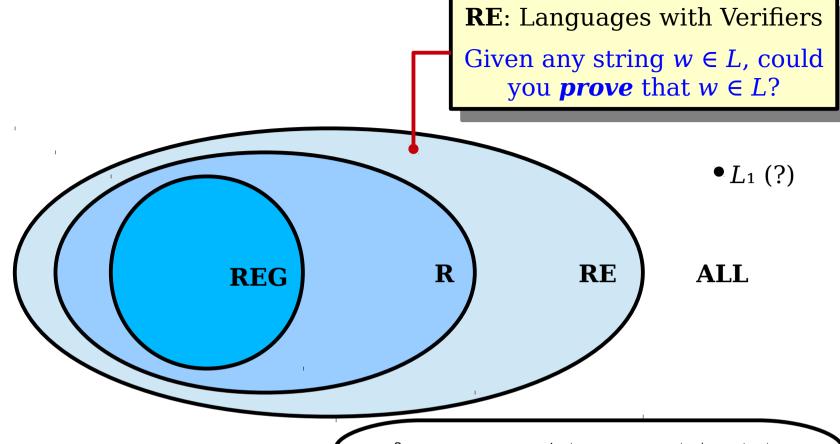


In other words, if we came across someone who was skeptical that the machine actually accepts at least two strings, could we convince them that the machine indeed does accept at least two strings?

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```







If we happened to know at least two strings that the machine accepted, we could just run the machine on both those strings and watch it accept them.

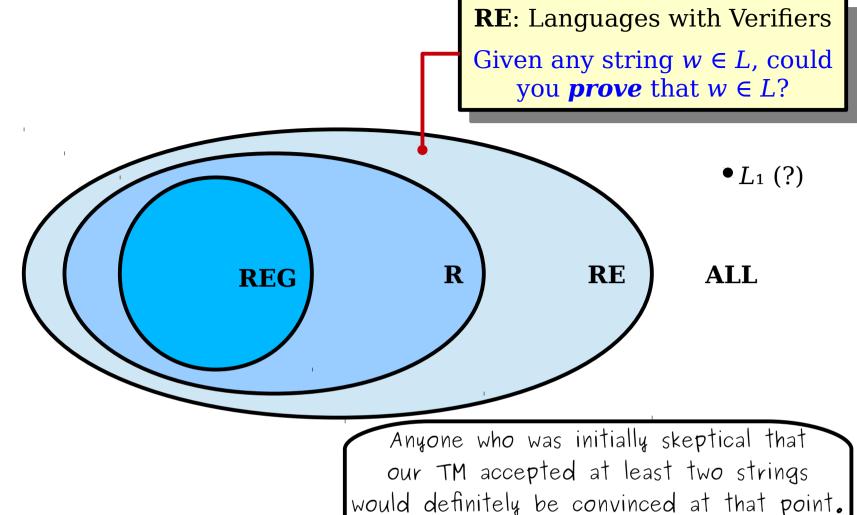
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}

L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

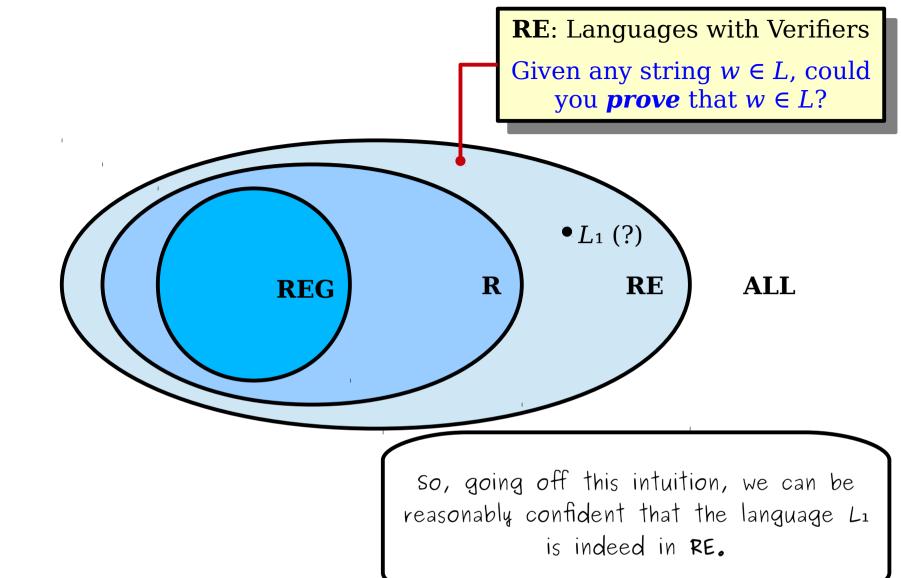




would definitely be convinced at that point. They just watched the TM accept at least two strings!

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

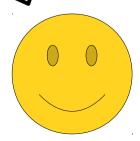


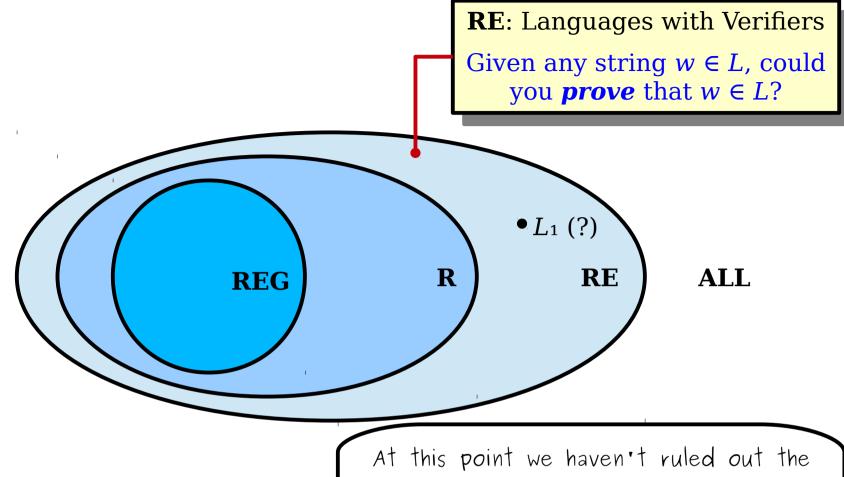


 $L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}$ $L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}$

 $L_3 = \{ a^n b^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}$

 $L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}$





At this point we haven't ruled out the possibility that it's also in R or is regular, but it's almost certainly not outside RE.

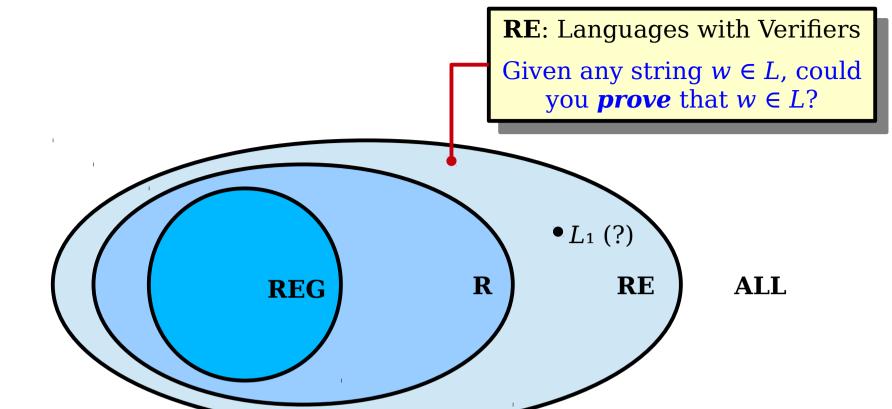
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}

L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

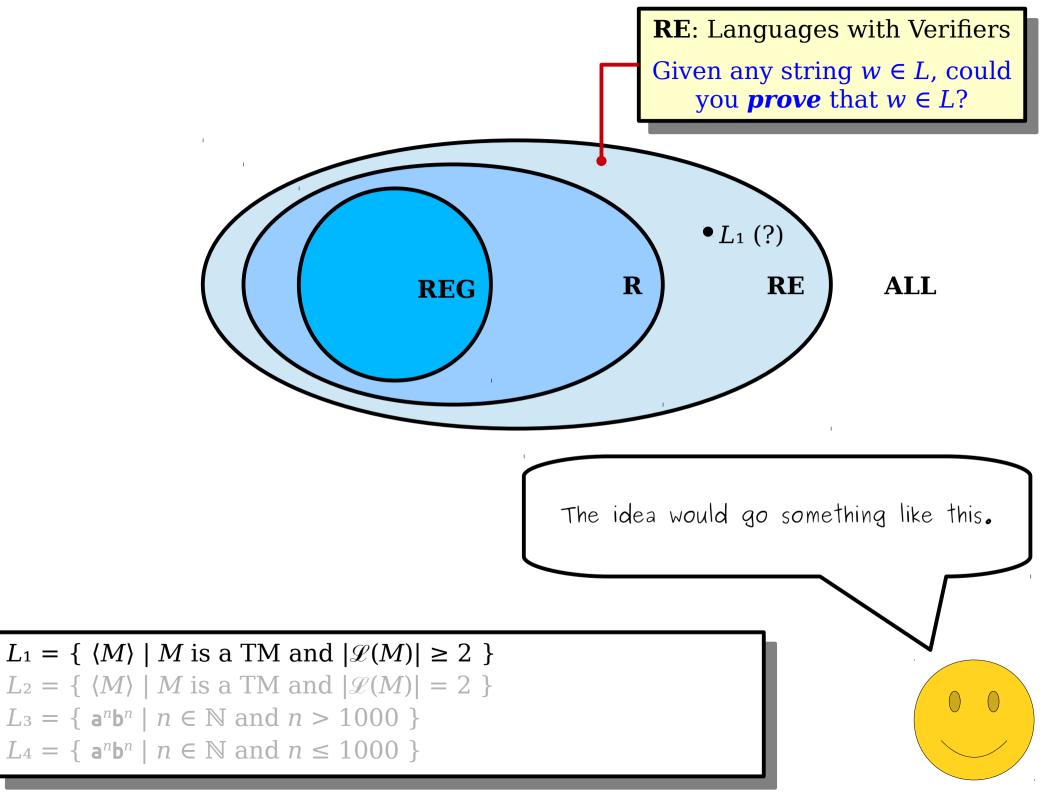


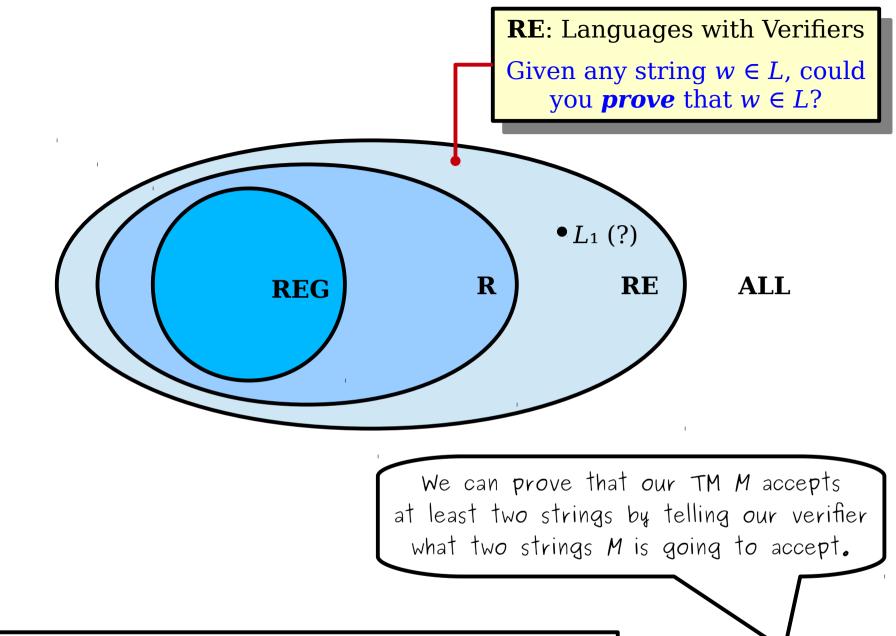


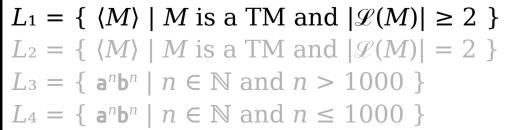
Although the question here was just to go and place L1, it's not a bad idea to think about how we'd actually go and build a verifier for L1.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

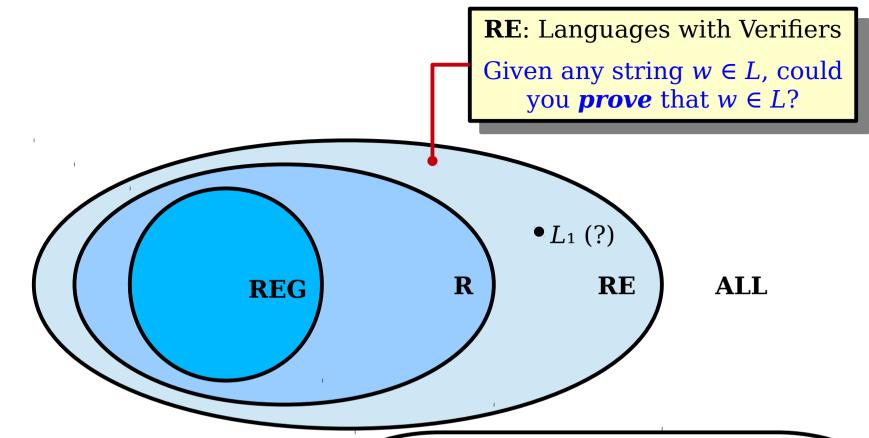








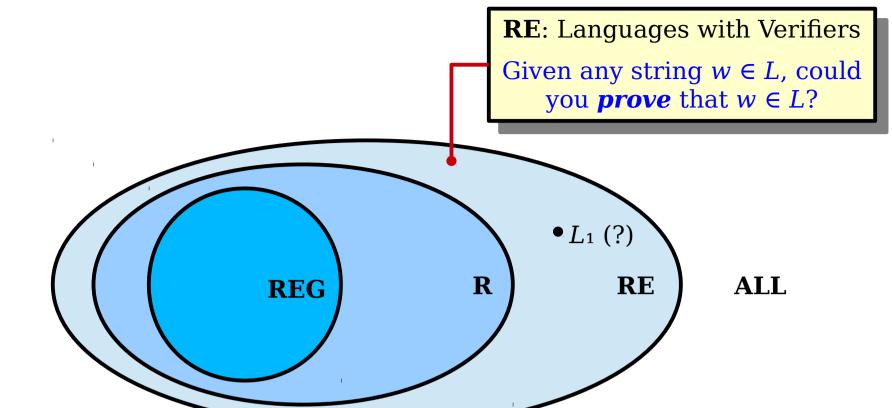




To ensure that our verifier doesn't go into an infinite loop (remember - verifiers aren't allowed to loop!), we can also give the verifier the number of steps it's going to take for M to accept.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

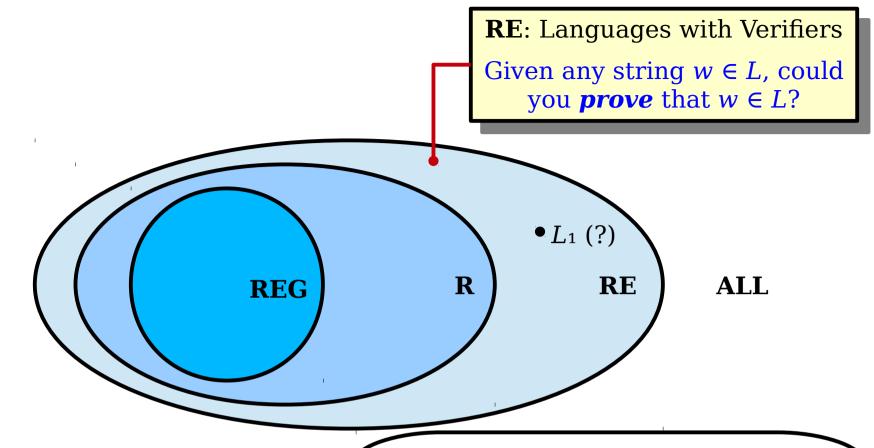




So the verifier would take in as input the TM M, two strings w_1 and w_2 , and a number of steps n, and could run M on the stings w_1 and w_2 for up to n steps.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

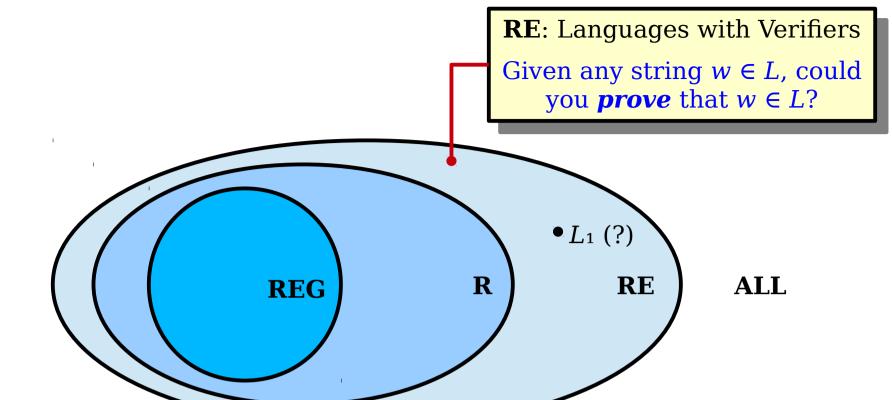




If M accepts both w1 and w2 within that many steps, then the verifier is convinced that M definitely accepts at least two strings.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

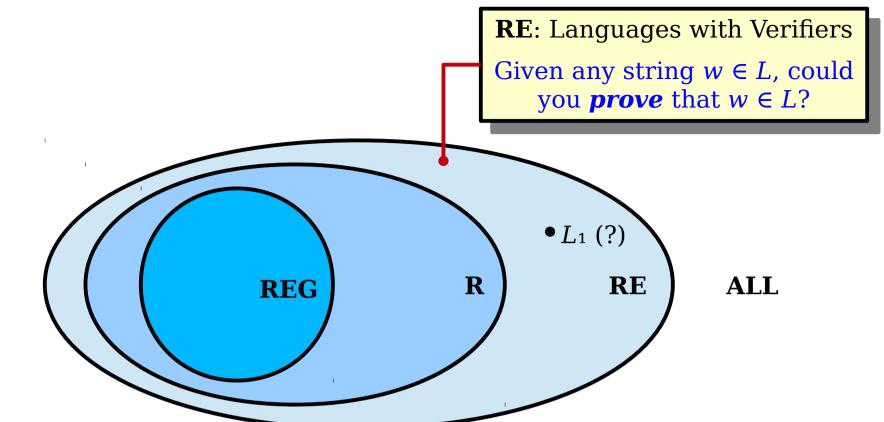




If that doesn't happen, the verifier isn't sure of what the answer is. Maybe M does accept two strings and we gave the verifier the wrong strings, or maybe we gave it the wrong number of steps.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

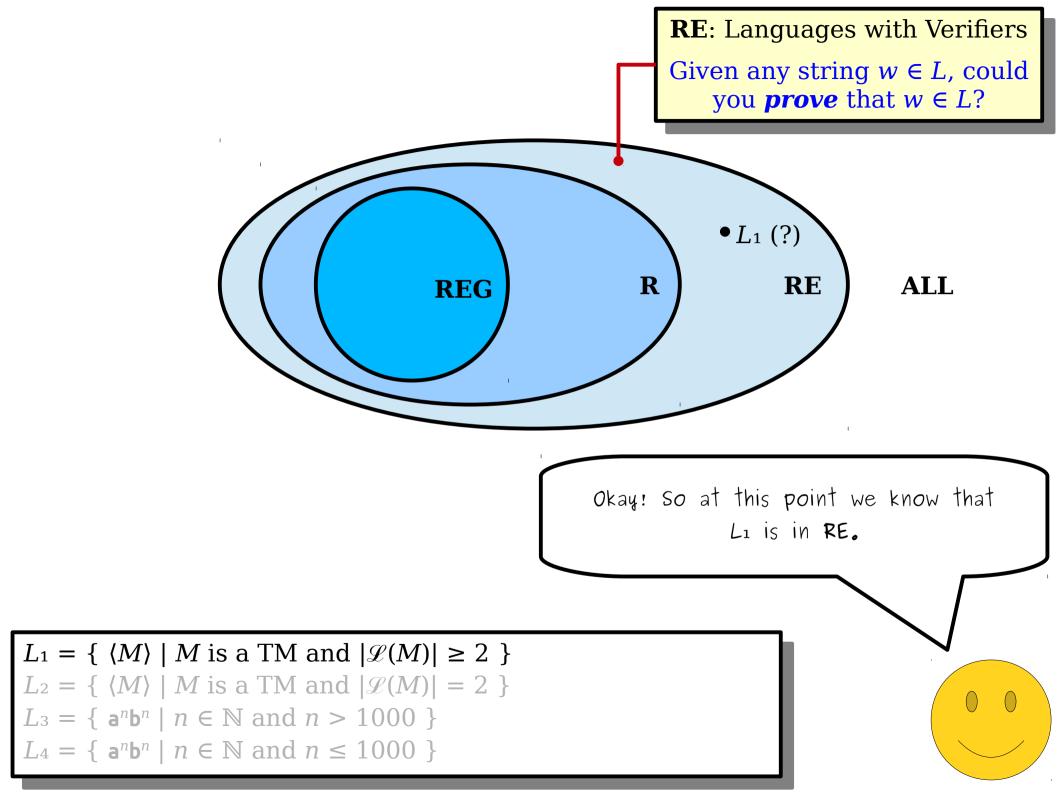


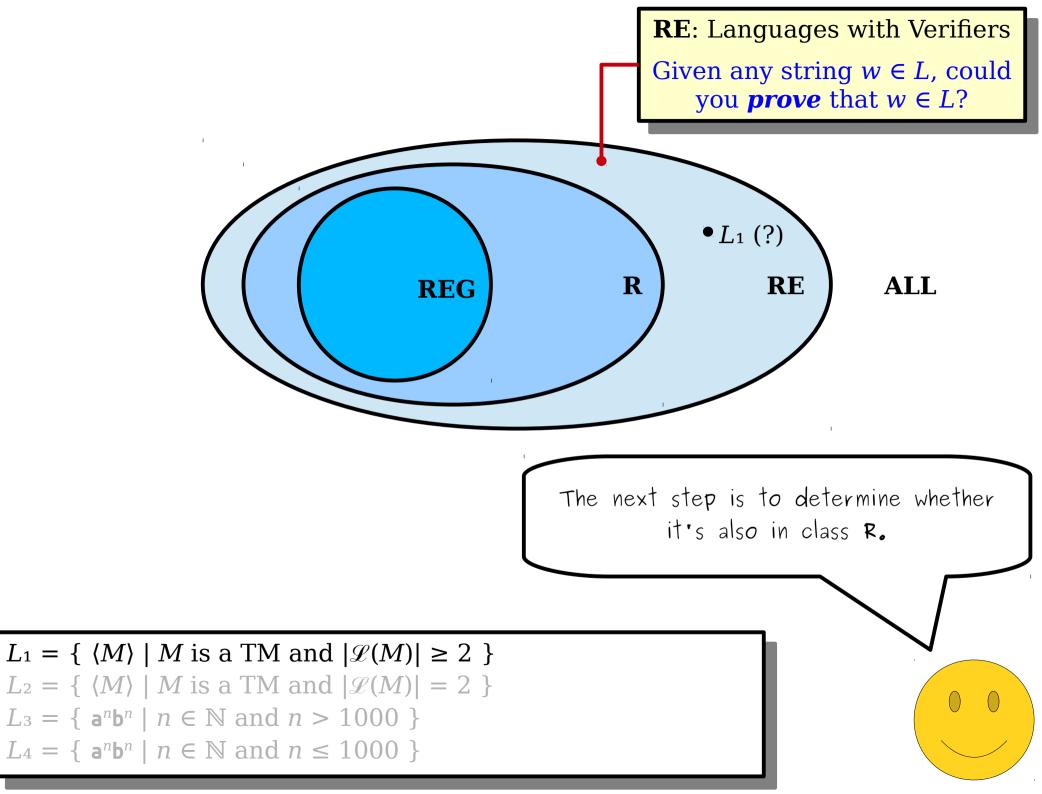


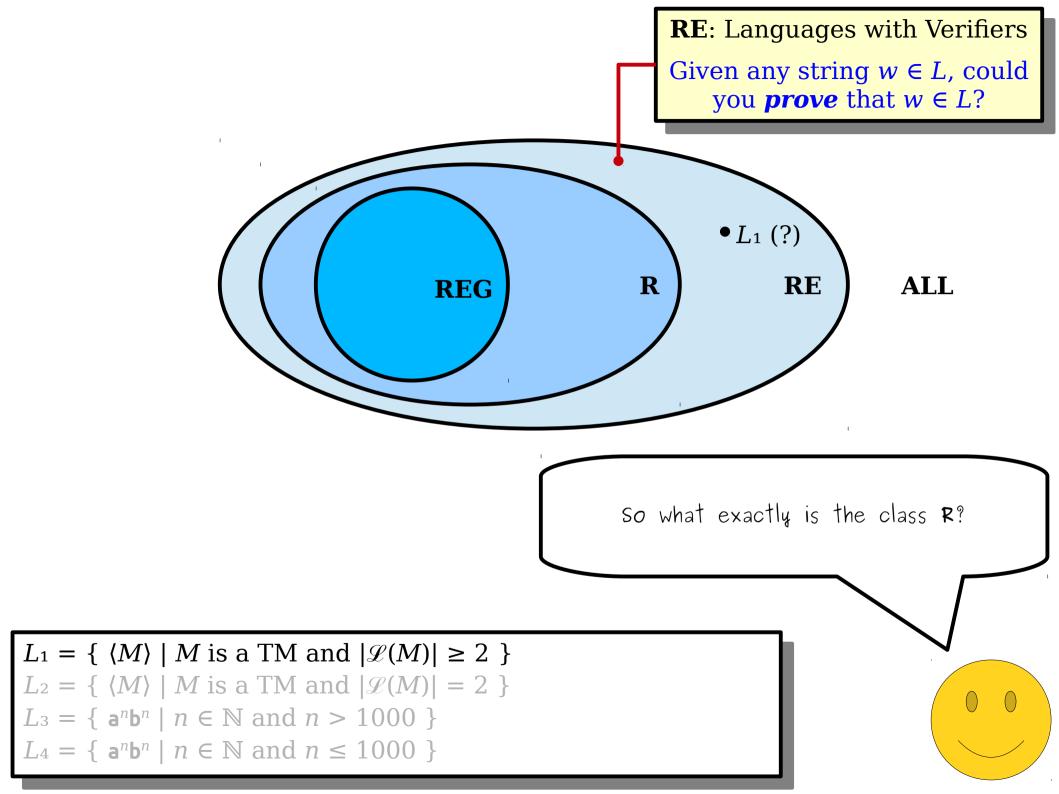
If you wanted to write this up as a formal proof, it's a good exercise! For now, though, we're just going to continue working through figuring out where this language goes on the Lava Diagram.

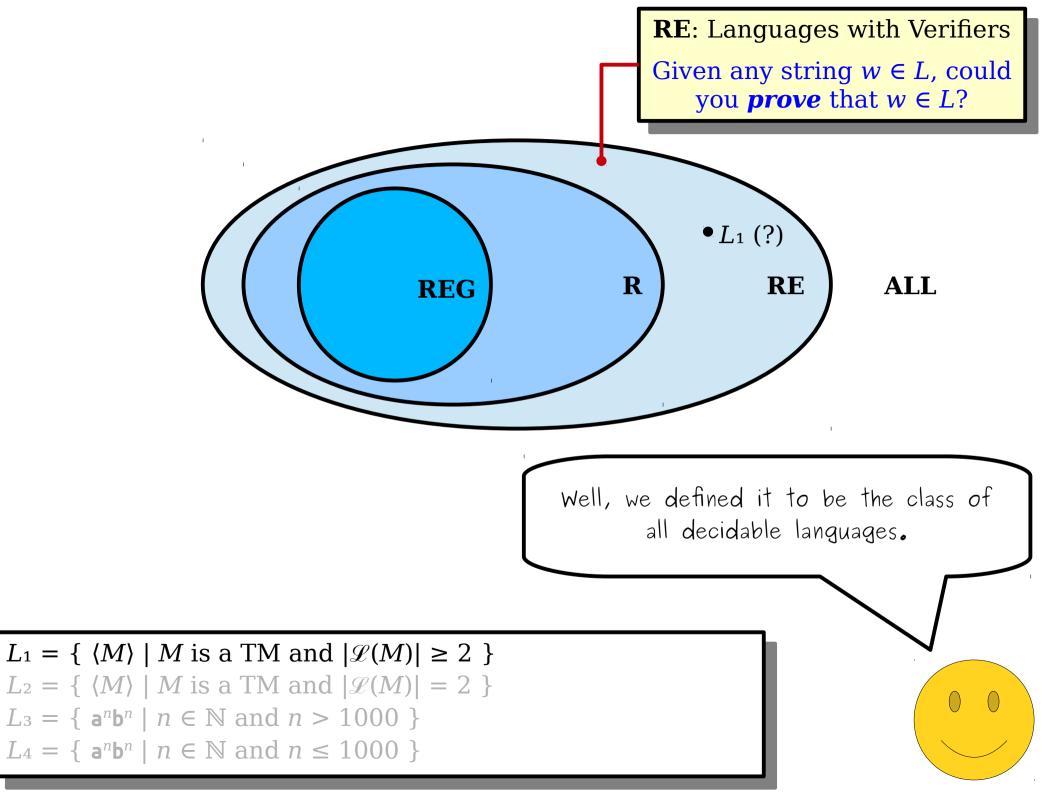
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

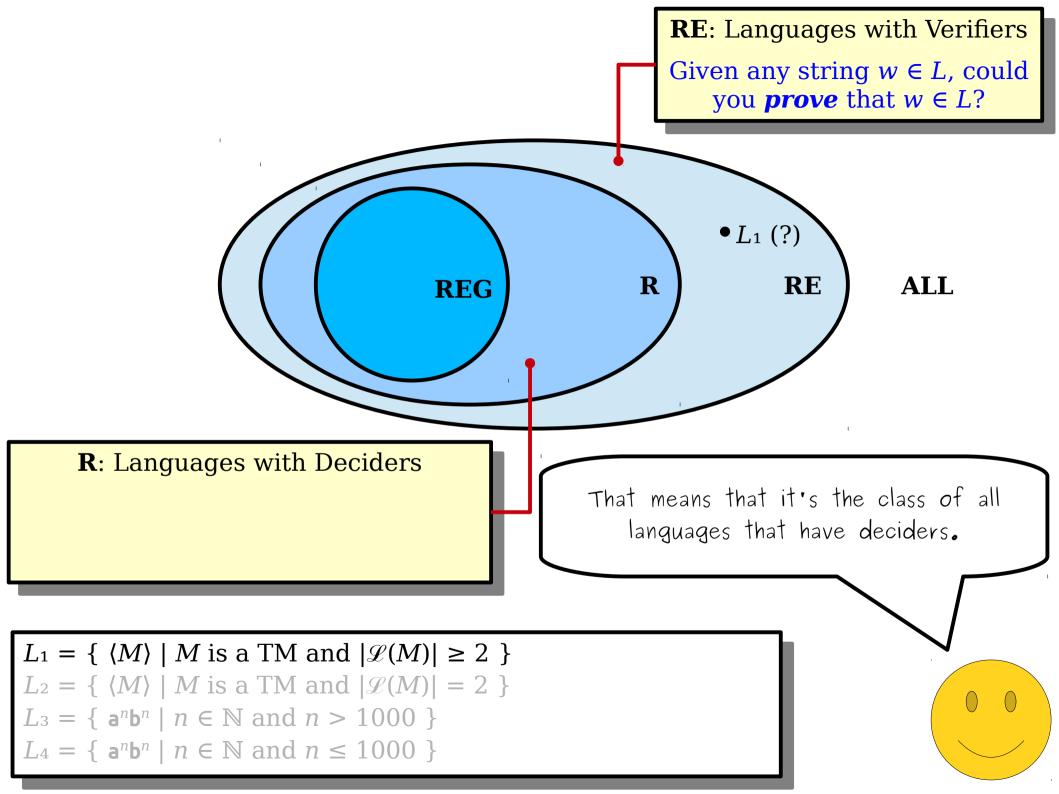


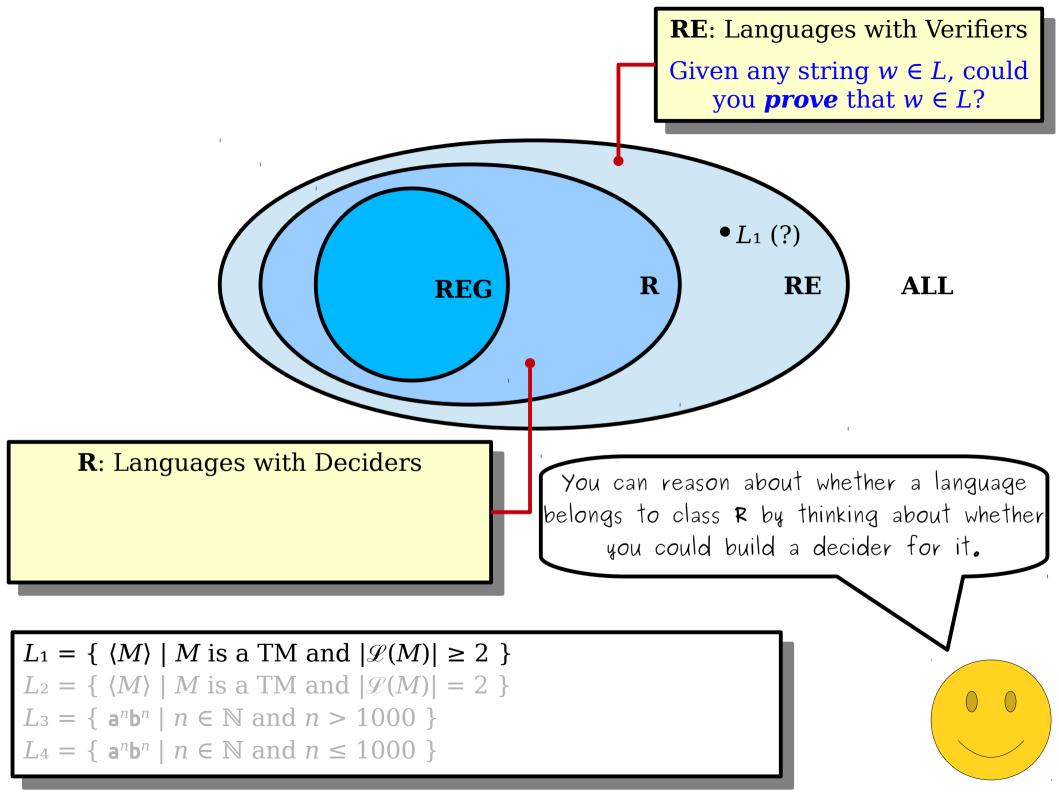


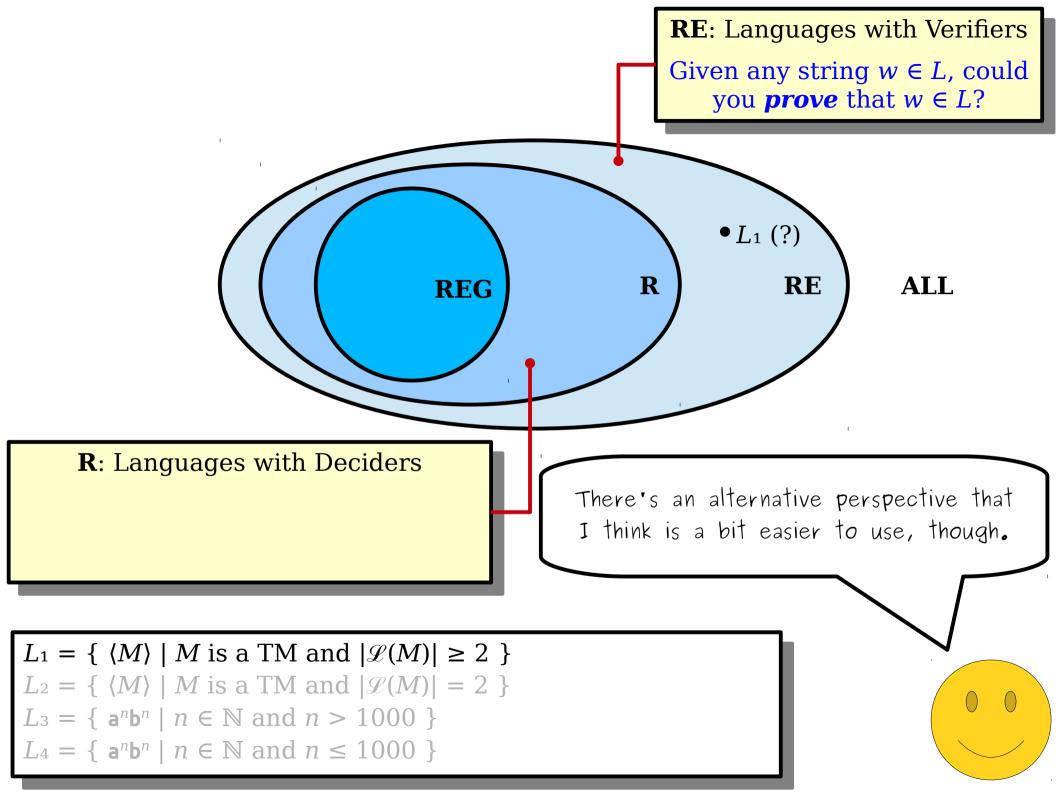


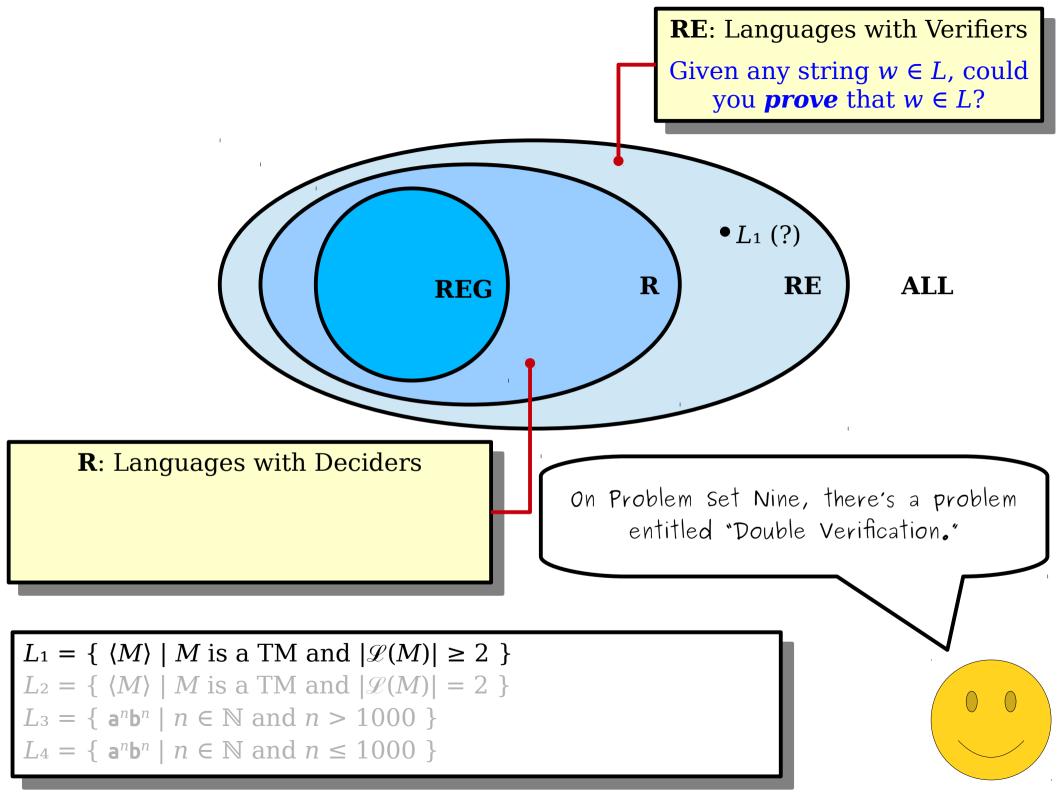


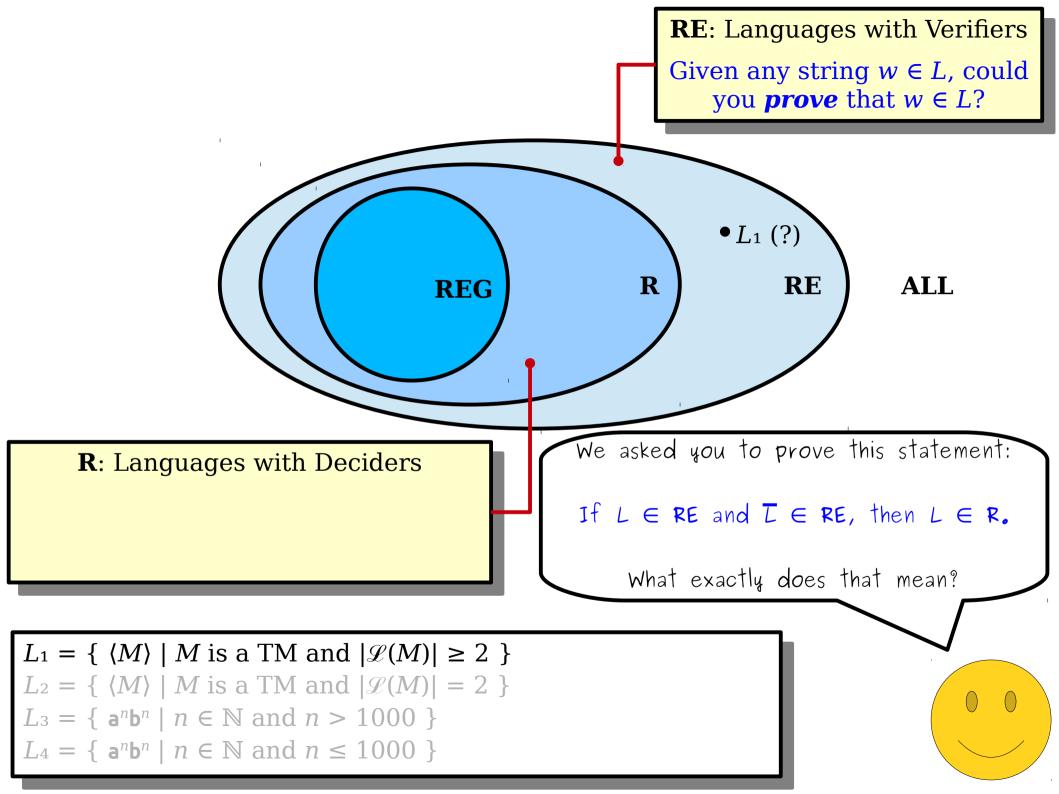


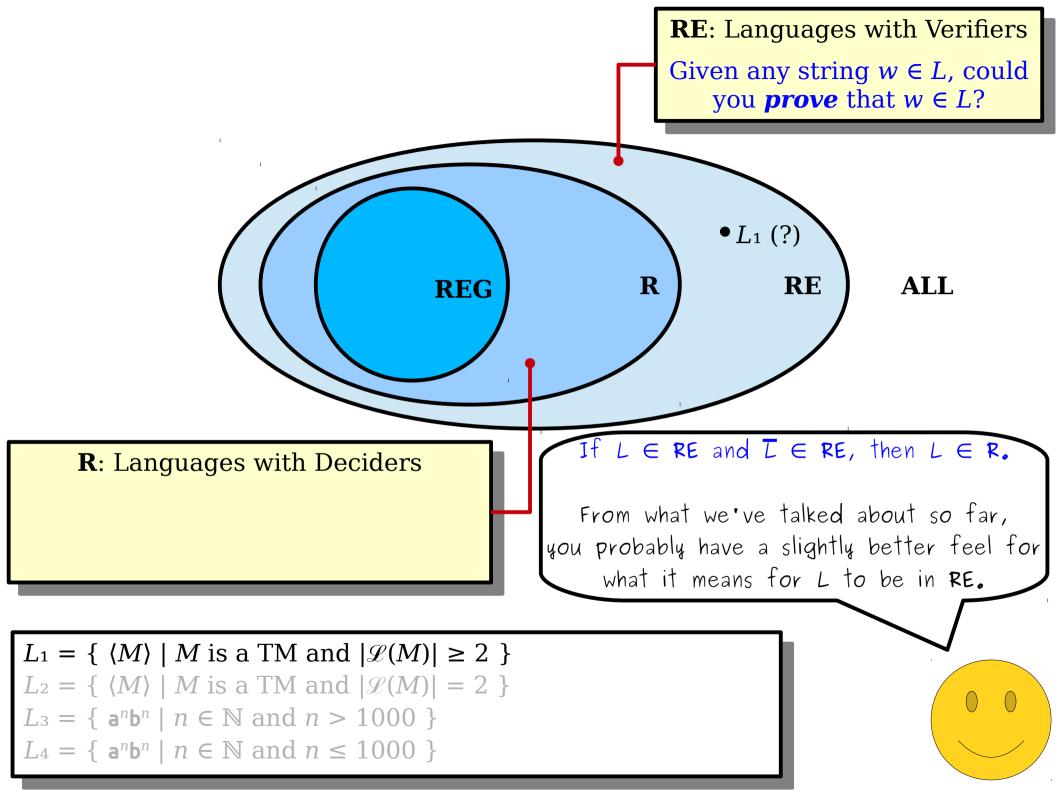


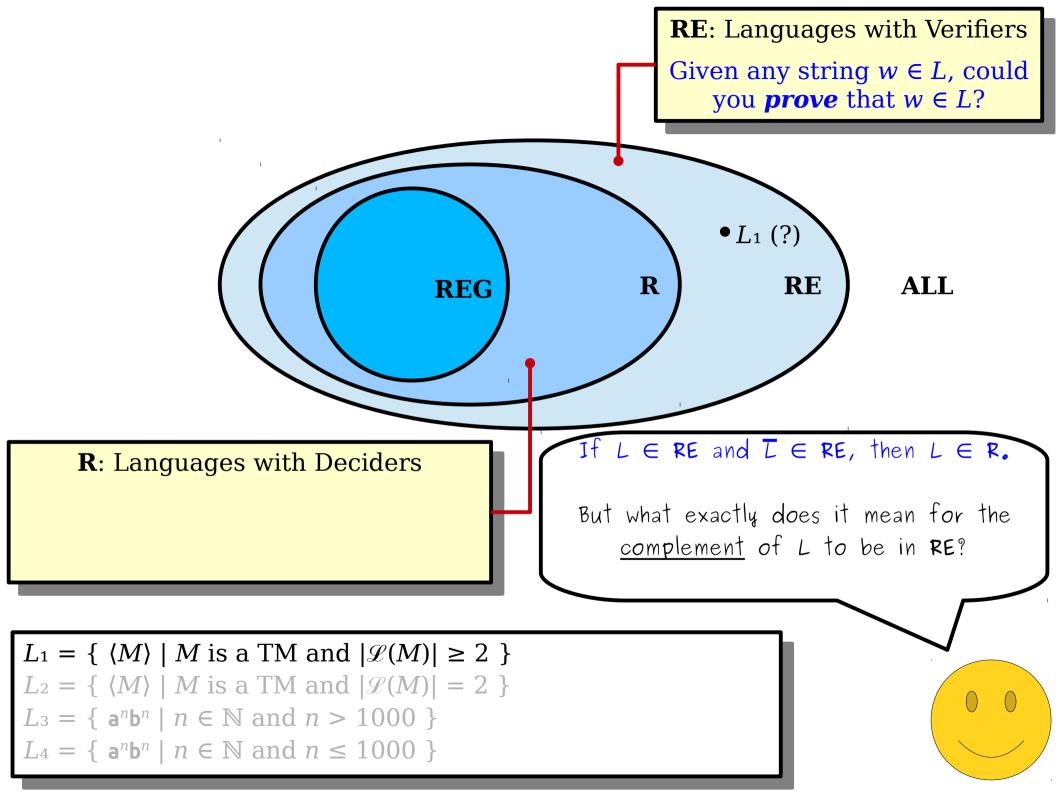


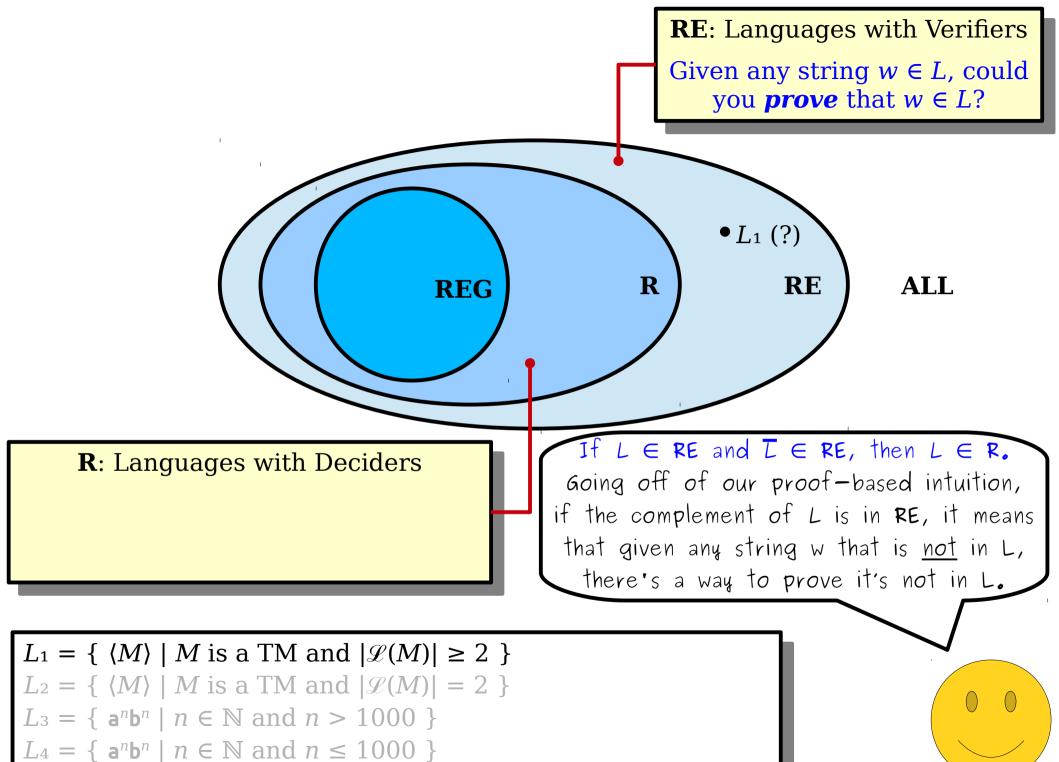


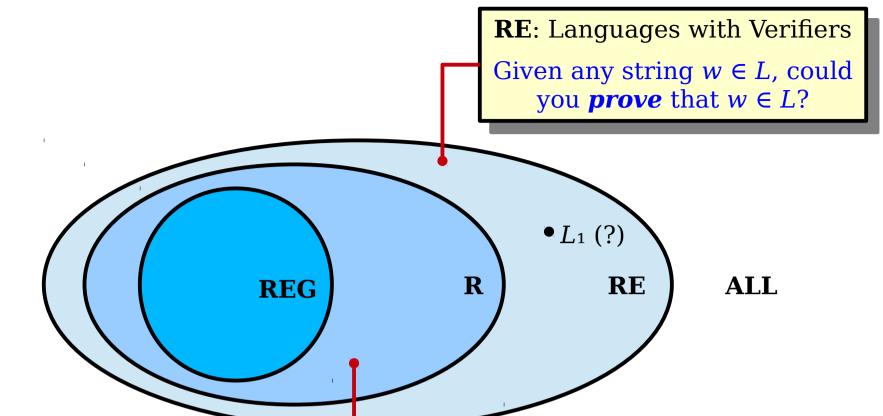












In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

This turns out to be a great way of intuiting the class R. A language belongs to R if it's in RE, and for any string that isn't in the language, there's a way to prove it's not in the language.

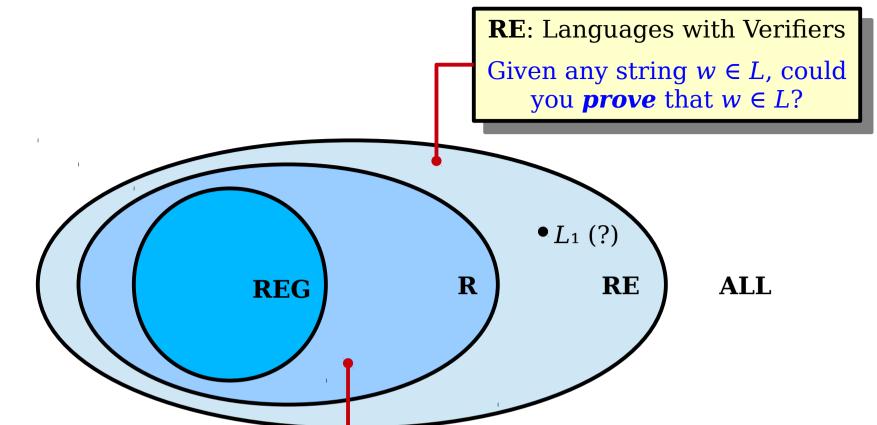
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}

L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

(Although we only had you prove the forward direction of the implication in the Double Verification problem, turns out the reverse direction holds as well. This gives an exact characterization of R!)

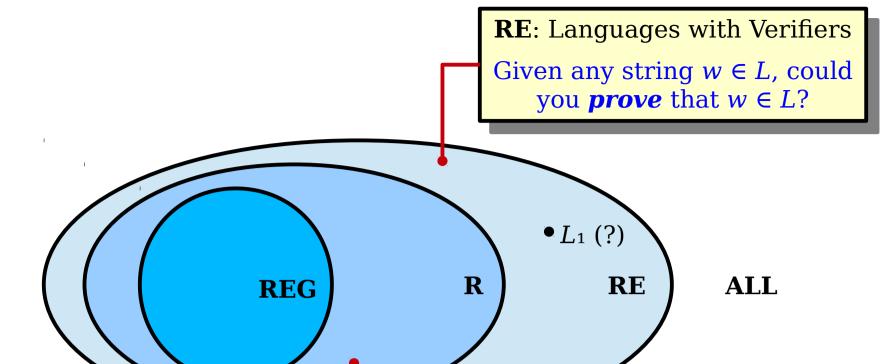
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}

L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```



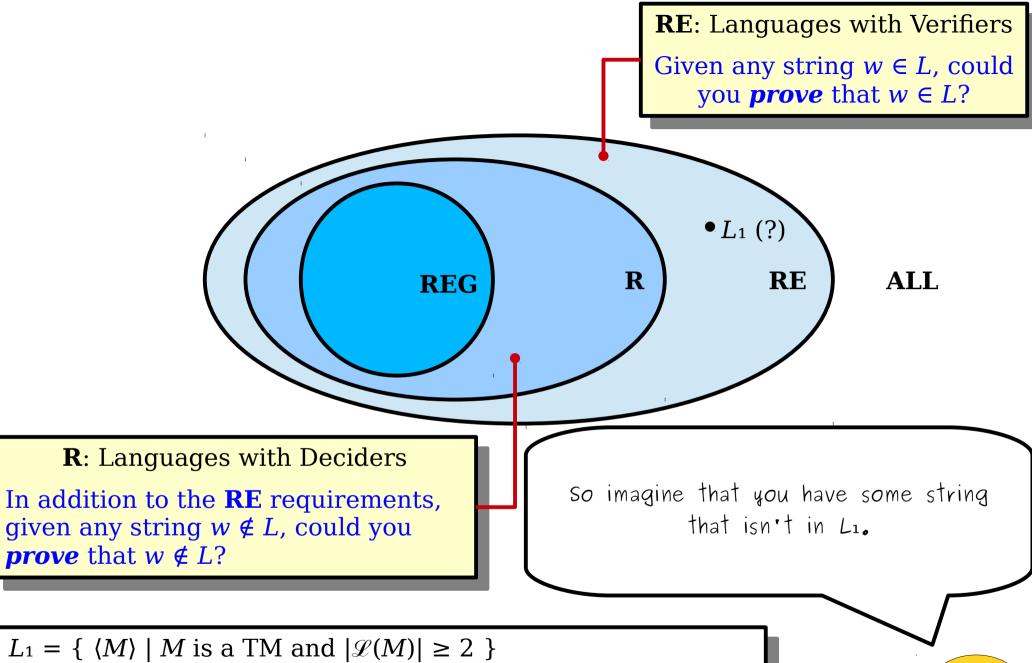


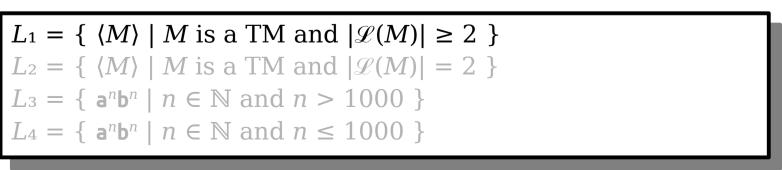
In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

Now, let's jump back to our particular language L1 here and use this intuition to think about whether or not it belongs to class R.

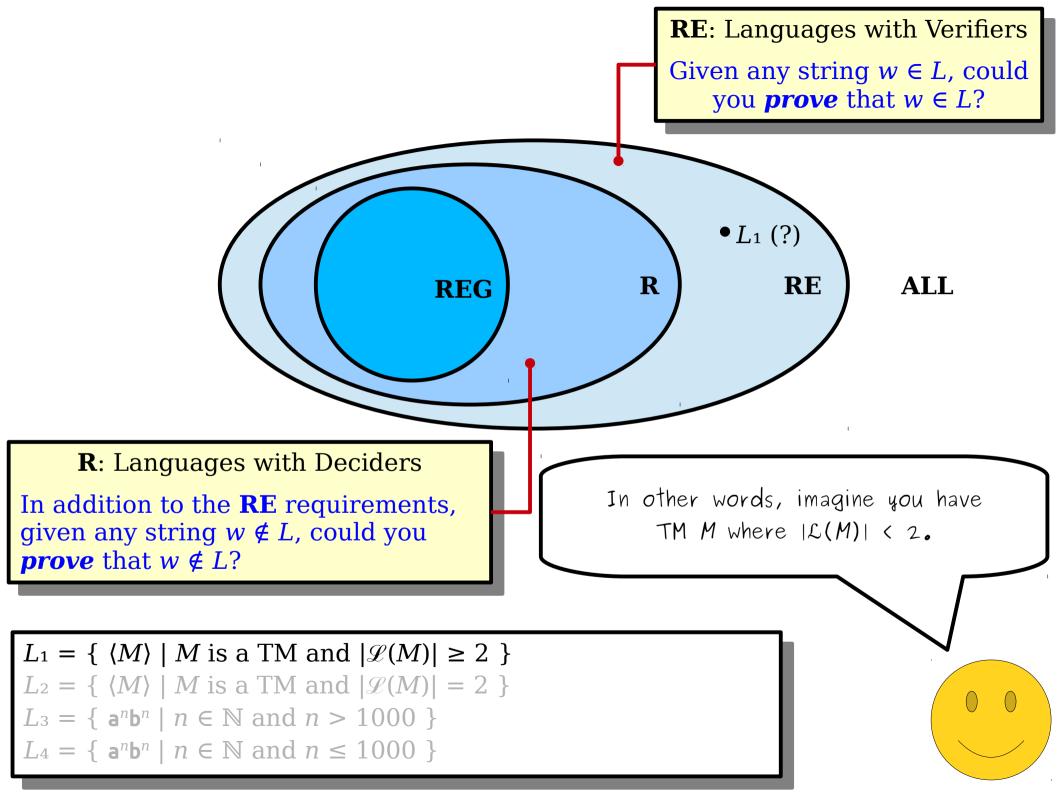
$$L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}$$
 $L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}$
 $L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}$
 $L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}$

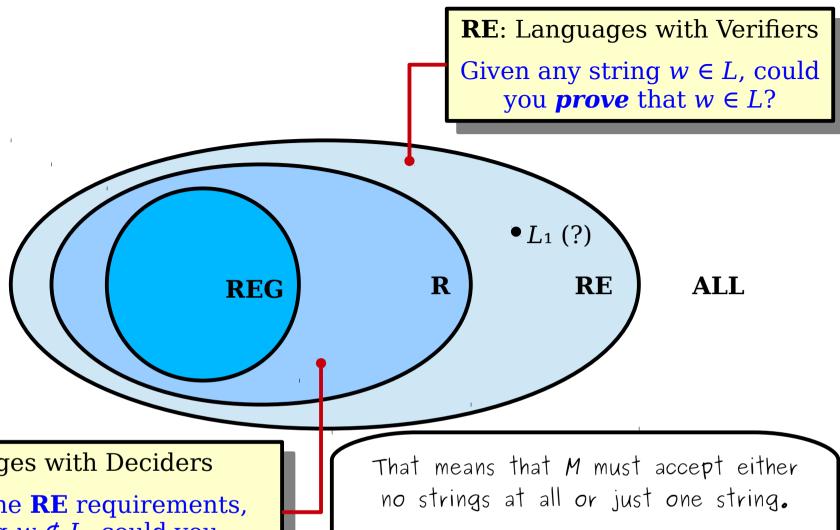










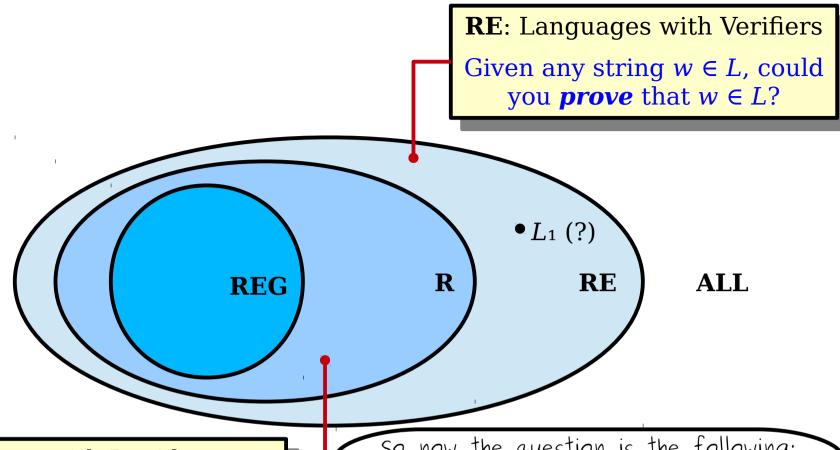


In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

(Do you see why?)

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ a^n b^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

So now the question is the following: if you have a TM that accepts either no strings or just one string, could you prove it to someone who was skeptical but honest?

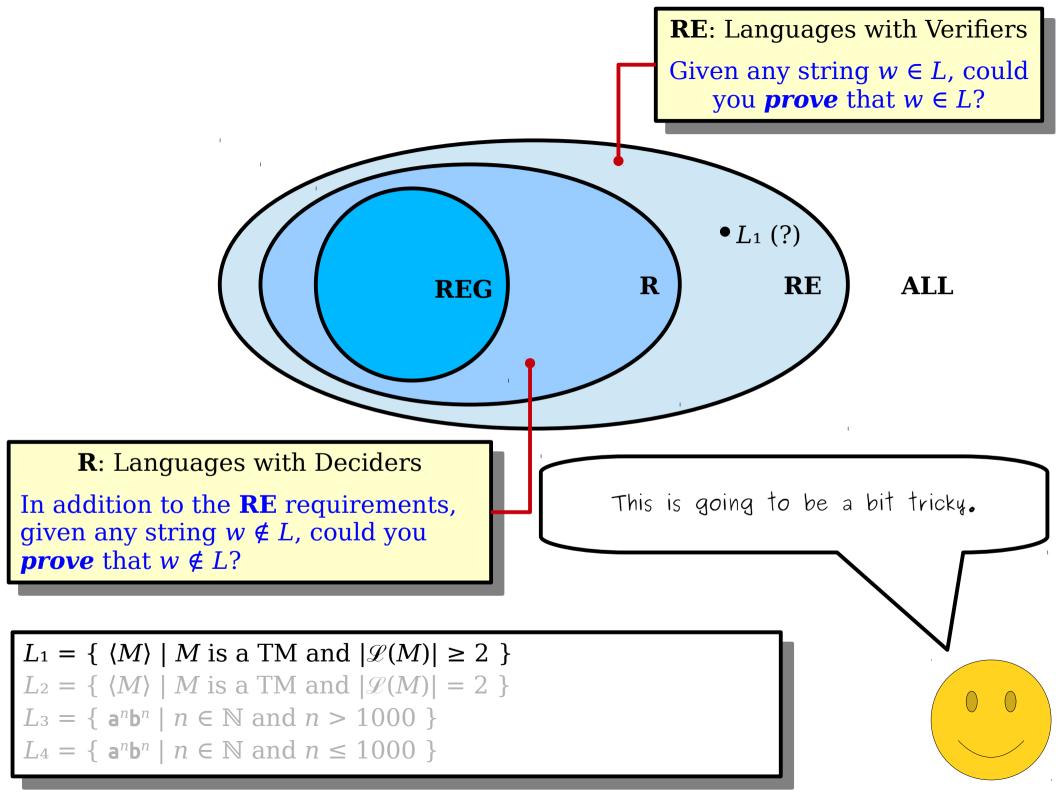
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}

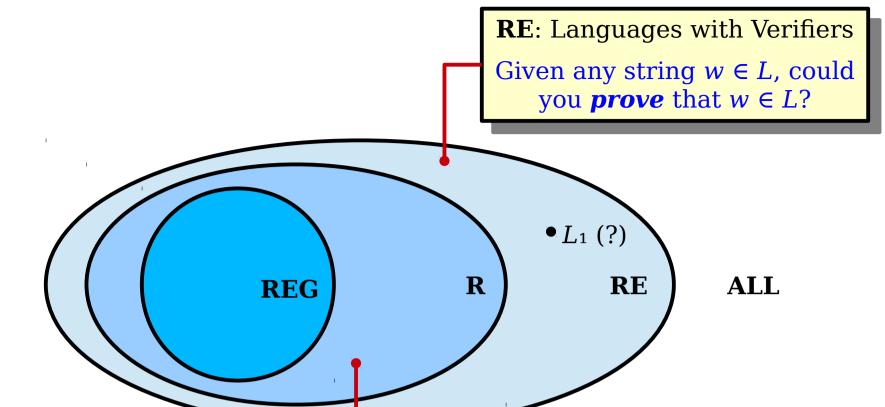
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```







In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

IF you want to convince someone that M only accepts at most one string, you need to convince them that out of the infinitely many strings that are out there, the TM accepts at most one.

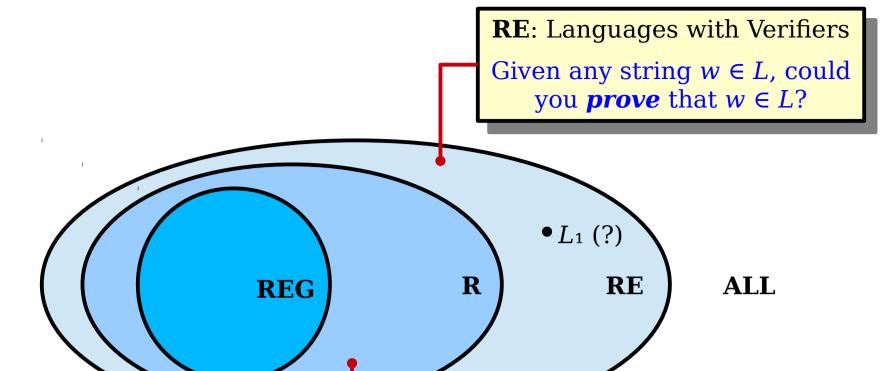
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}

L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





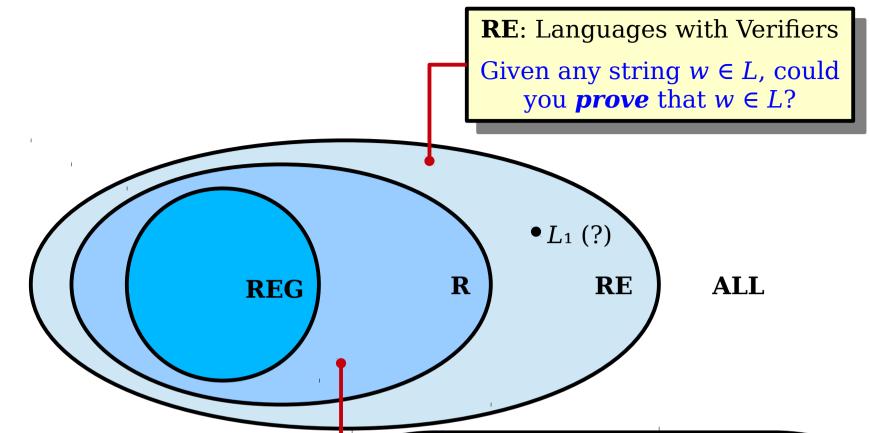
In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

As we've seen before, though, we know that the only general way to find out what a TM will do on a string is to run the TM on that string and see what happens.

$$L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}$$

 $L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}$
 $L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}$
 $L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}$





In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

So if we want to convince someone that a TM doesn't accept infinitely many different strings, we're out of luck! In the general case, we'd have to run the TM on all those strings...

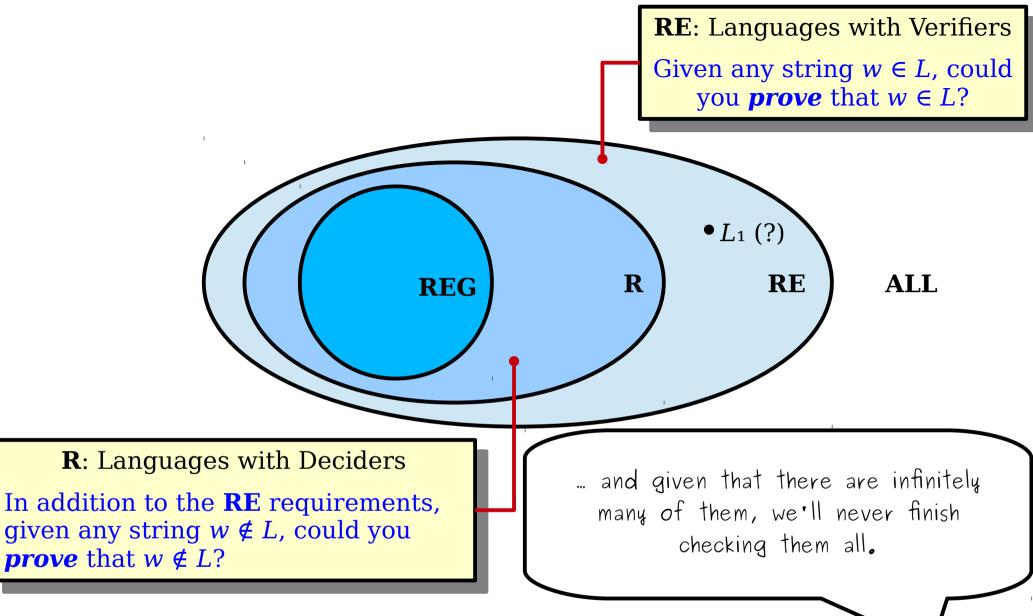
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}

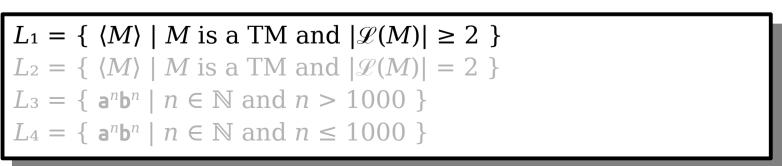
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

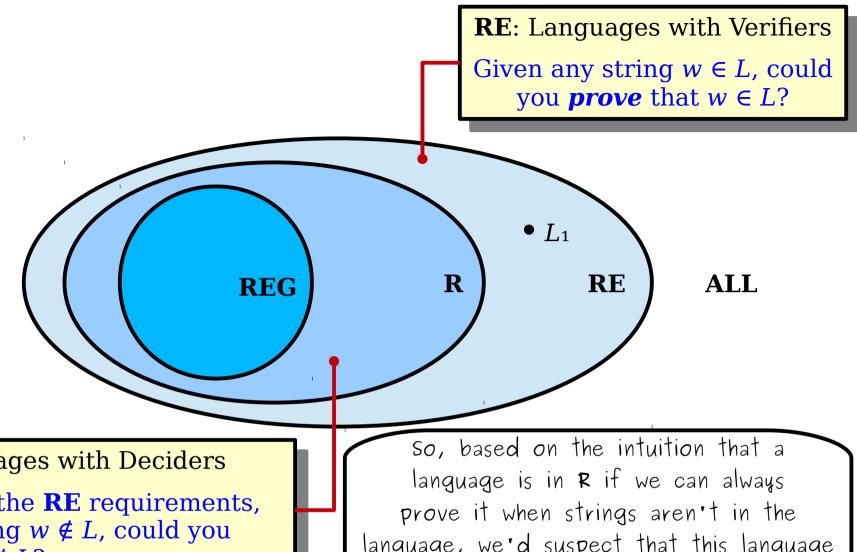
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```









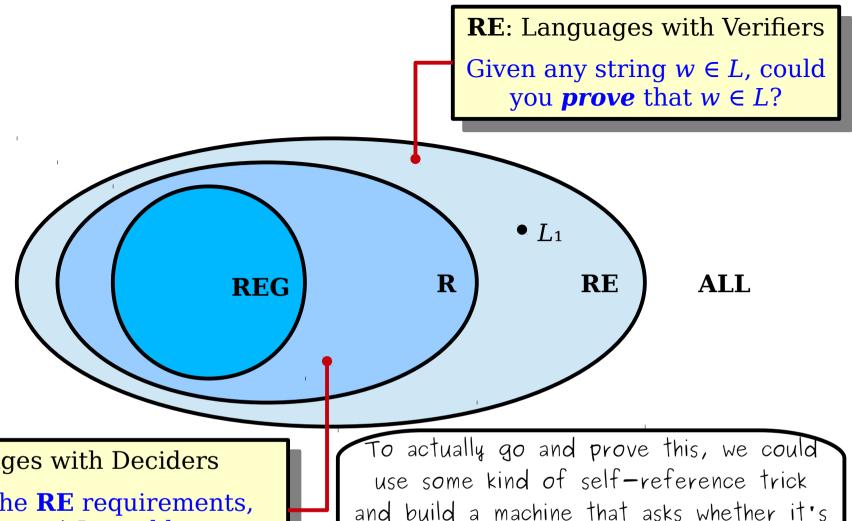


In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

language, we'd suspect that this language is not in R.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ a^n b^n \mid n \in \mathbb{N} \text{ and } n > 1000 \} 
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```



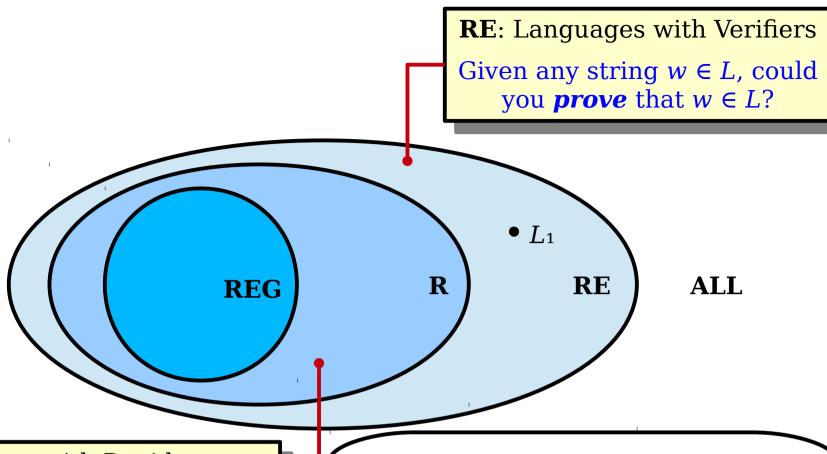


In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

and build a machine that asks whether it's going to accept at least two strings, then does the opposite.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ a^n b^n \mid n \in \mathbb{N} \text{ and } n > 1000 \} 
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





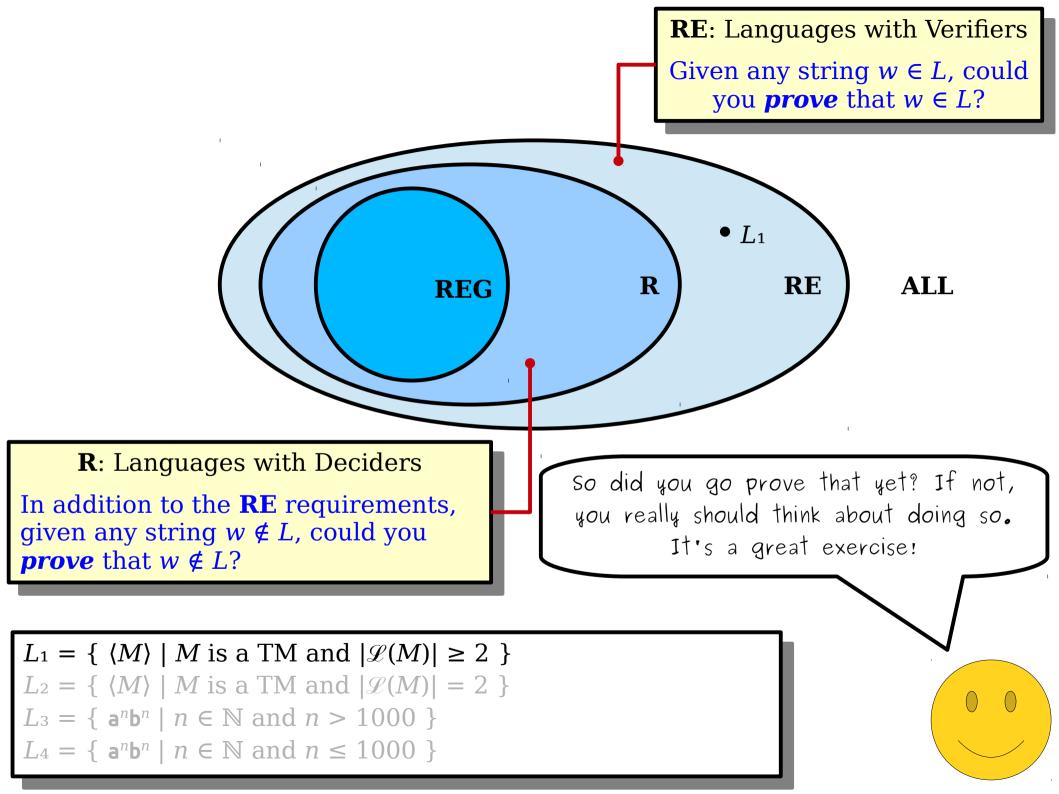
In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

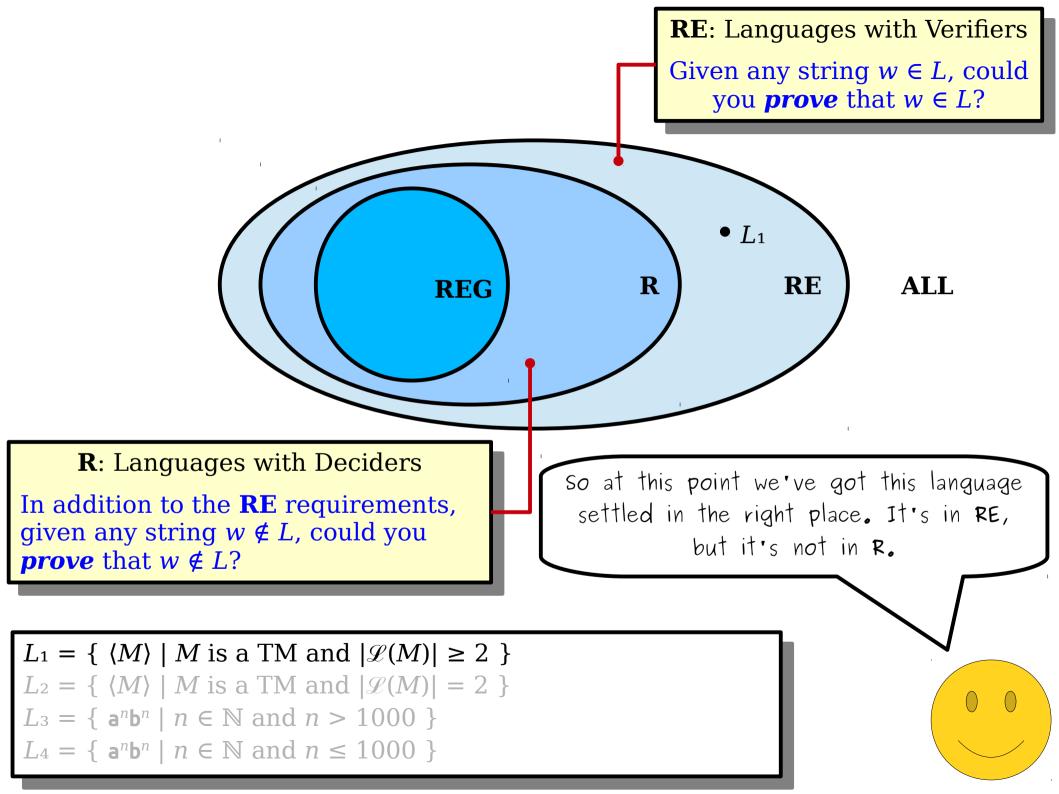
In fact, that's such a good exercise that you should stop reading this and go do it right now. The Guide to Self-Reference might help you there.

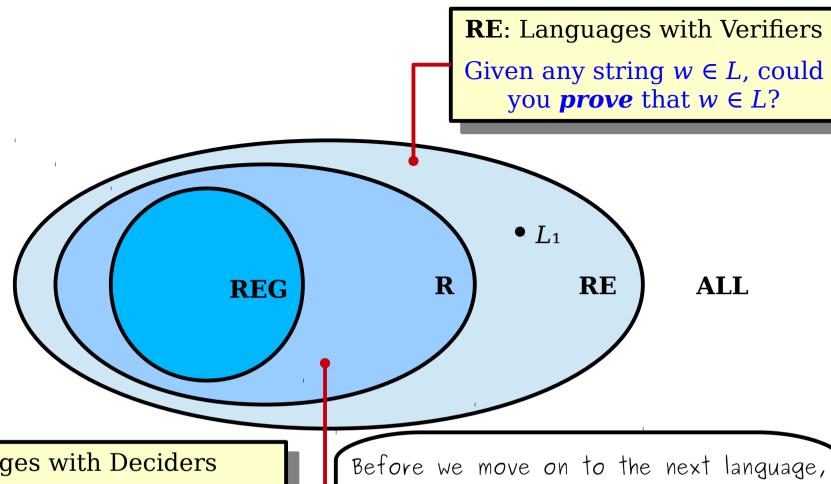
$$L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}$$

 $L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}$
 $L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}$
 $L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}$







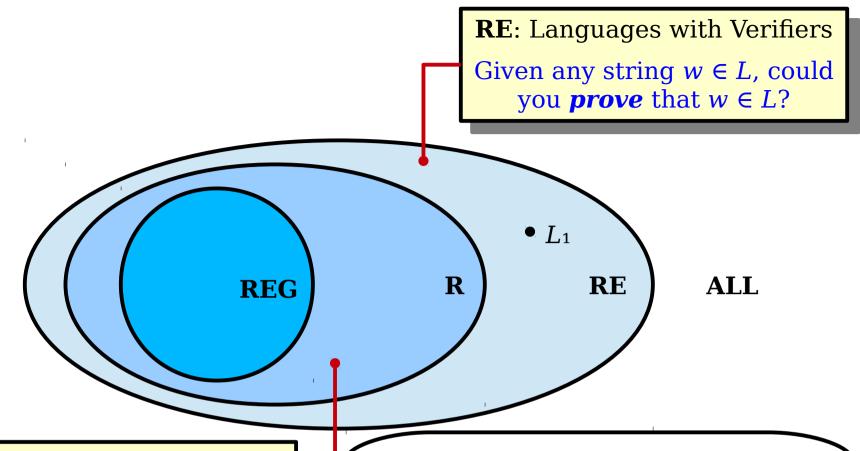


In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

I wanted to take a minute to address a common question we get on problems like these.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ a^n b^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```



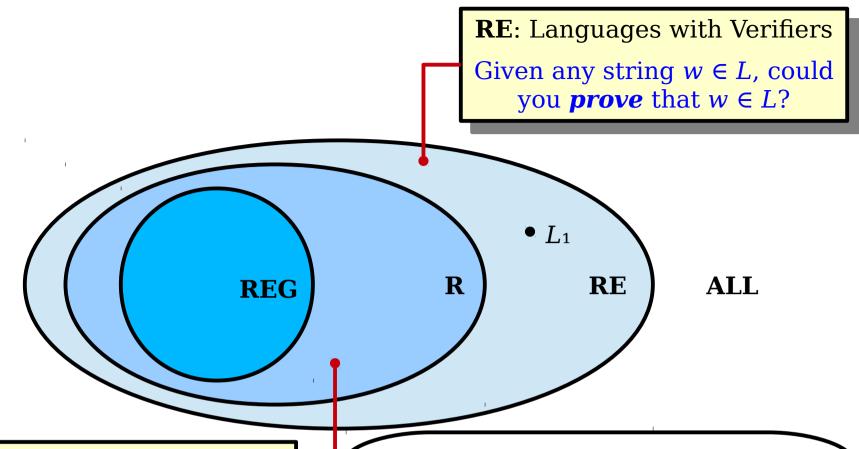


In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

If you look at the description of the language, you can see that it says something about TMs that accept at least two strings.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```



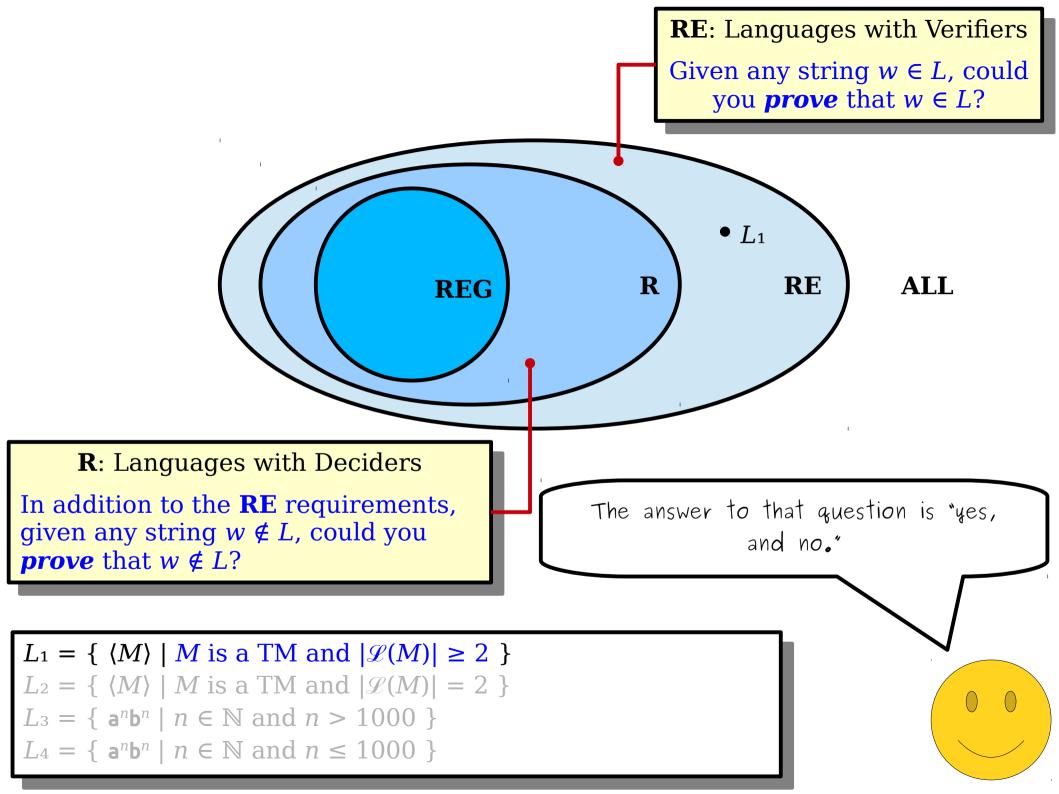


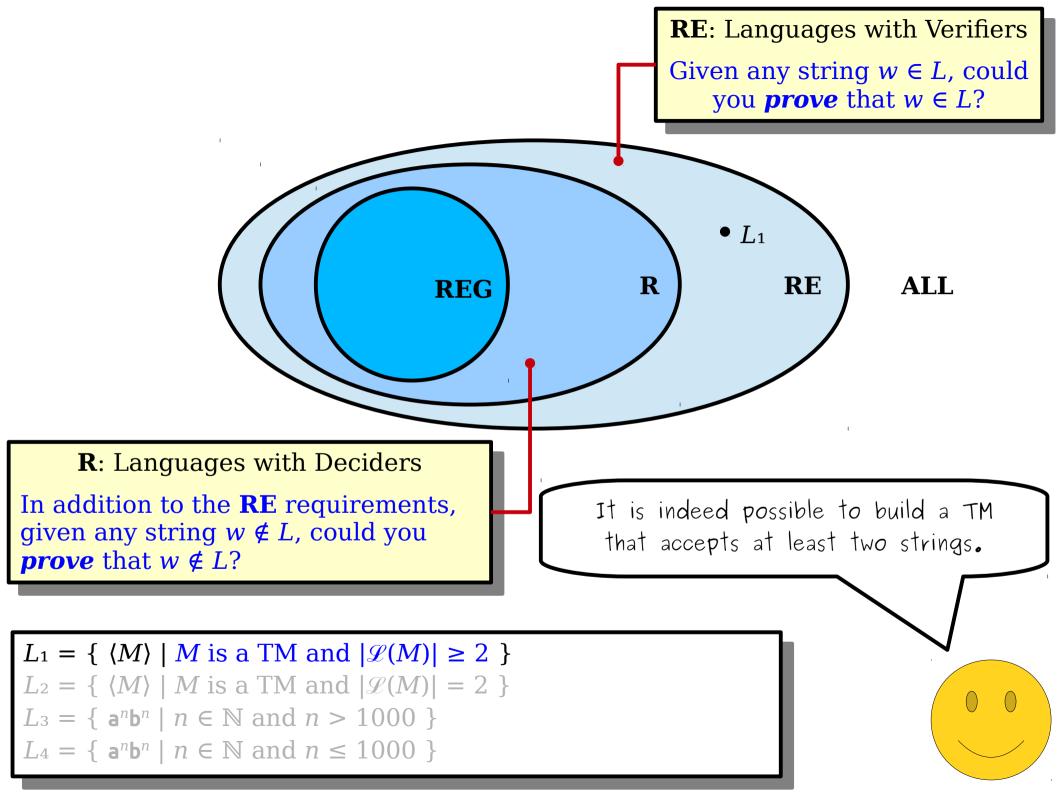
In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

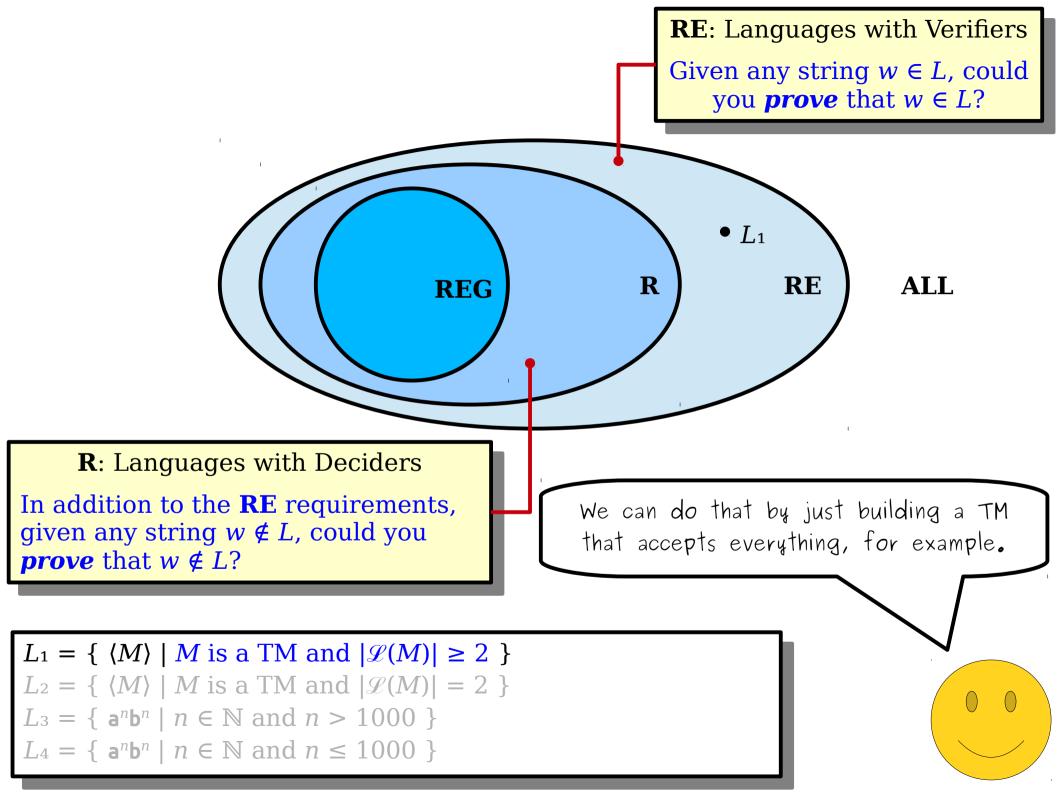
A lot of people ask - "Isn't it really easy to build a TM that accepts at least two strings? So shouldn't this be decidable? Or even regular?"

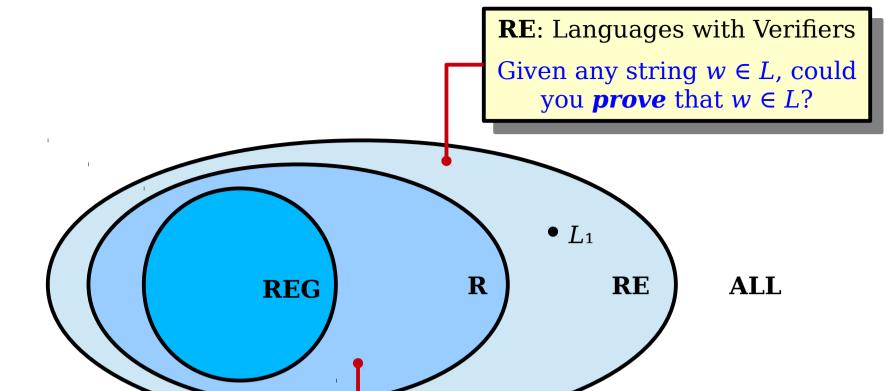
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```











In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

But notice that this problem isn't asking whether you can build this machine. It's a question about the language of all TMs with this particular property.

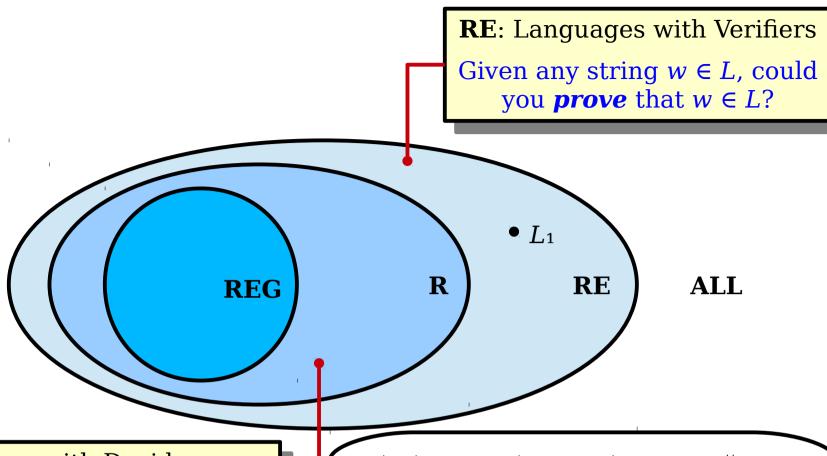
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}

L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```



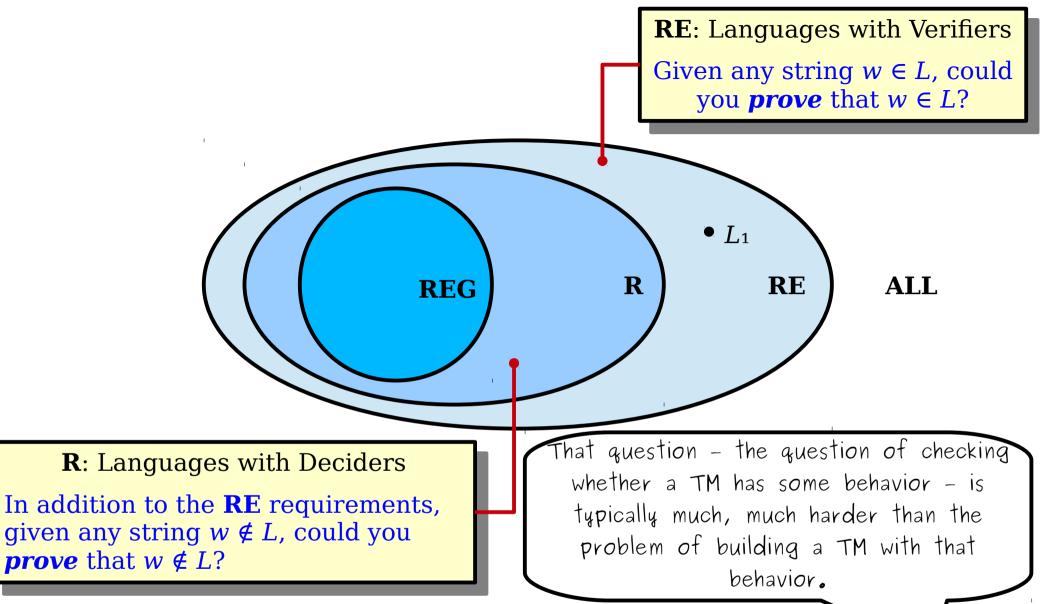


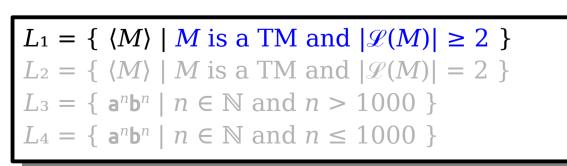
In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

In that sense, the question is really asking "how hard is it to tell whether a random TM actually does accept at least two strings?"

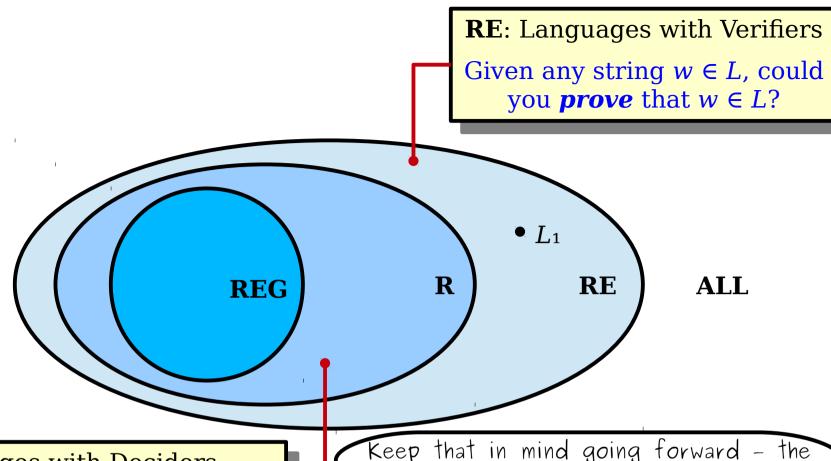
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```









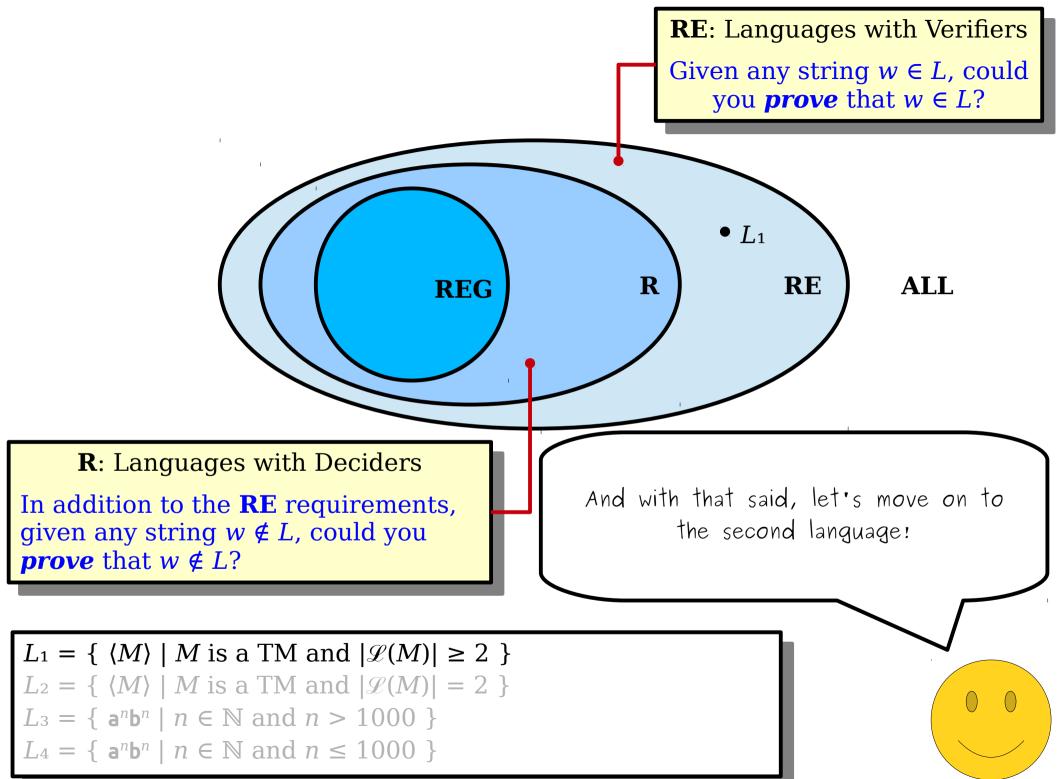


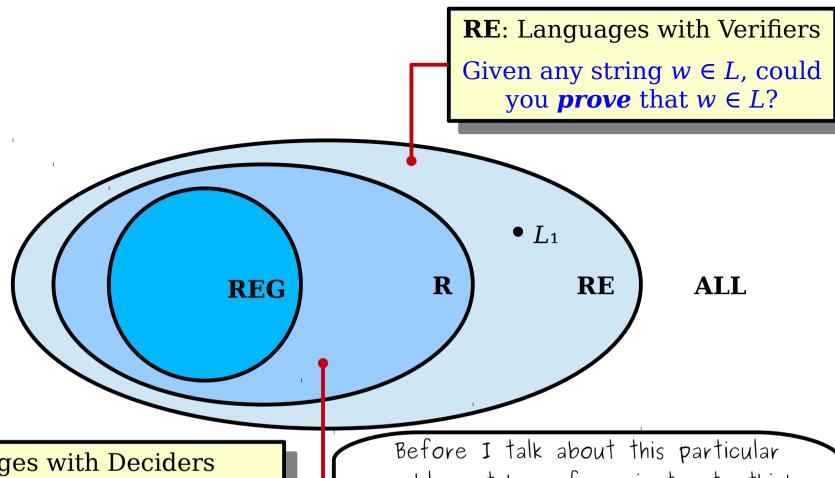
In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

Keep that in mind going forward - the question is to determine whether an arbitrary string is in the language, not to try to find a string that happens to be in the language.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





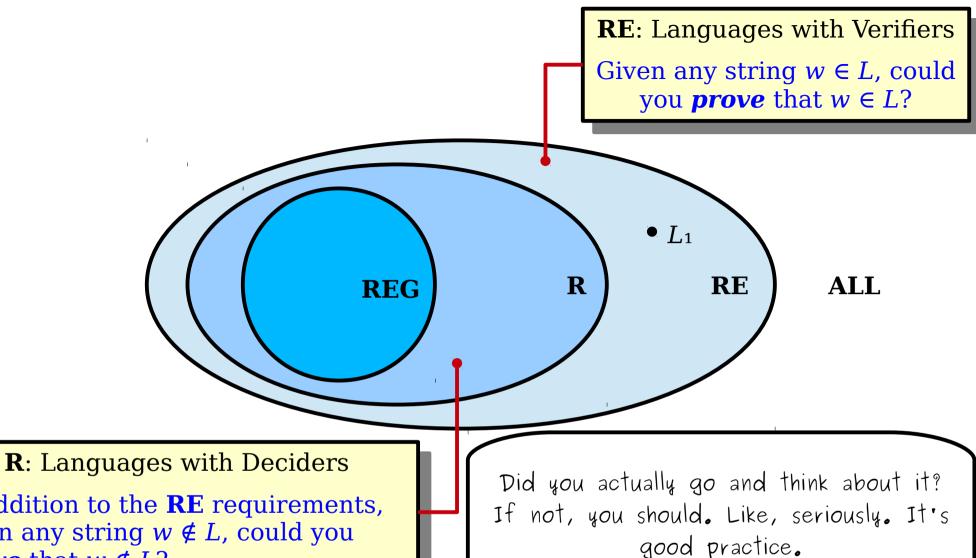


In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

problem, take a few minutes to think about where you believe this should go in the Lava Diagram. Once you've done that, let's rejoin and keep talking.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ a^n b^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

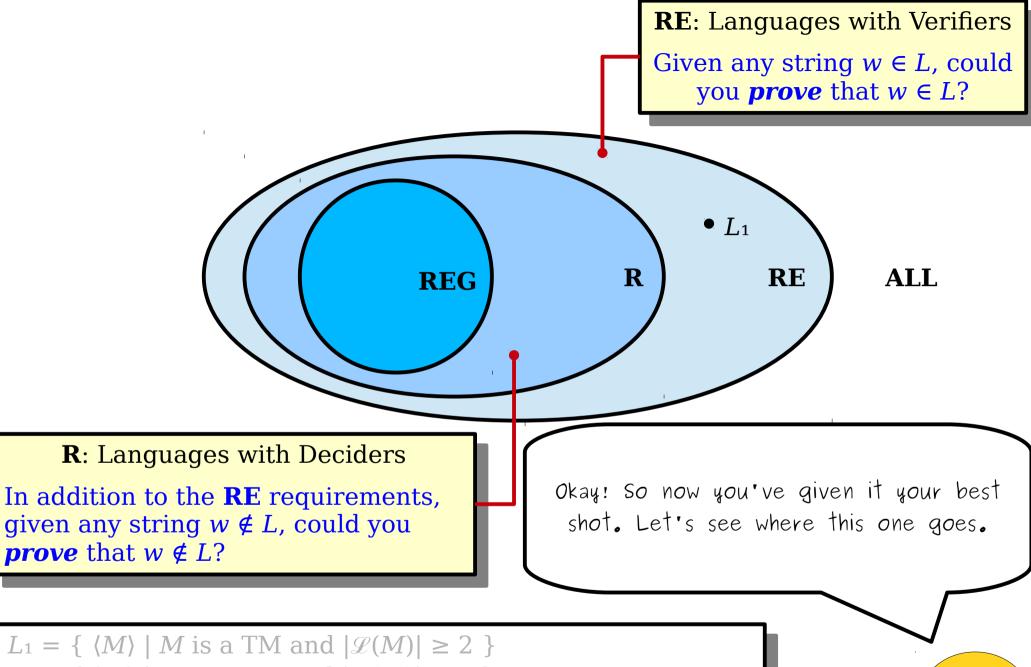




In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

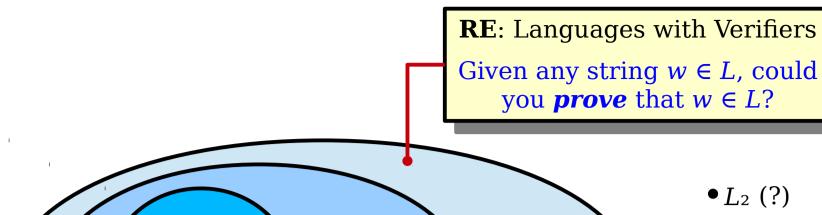
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ a^n b^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





REG R RE ALL

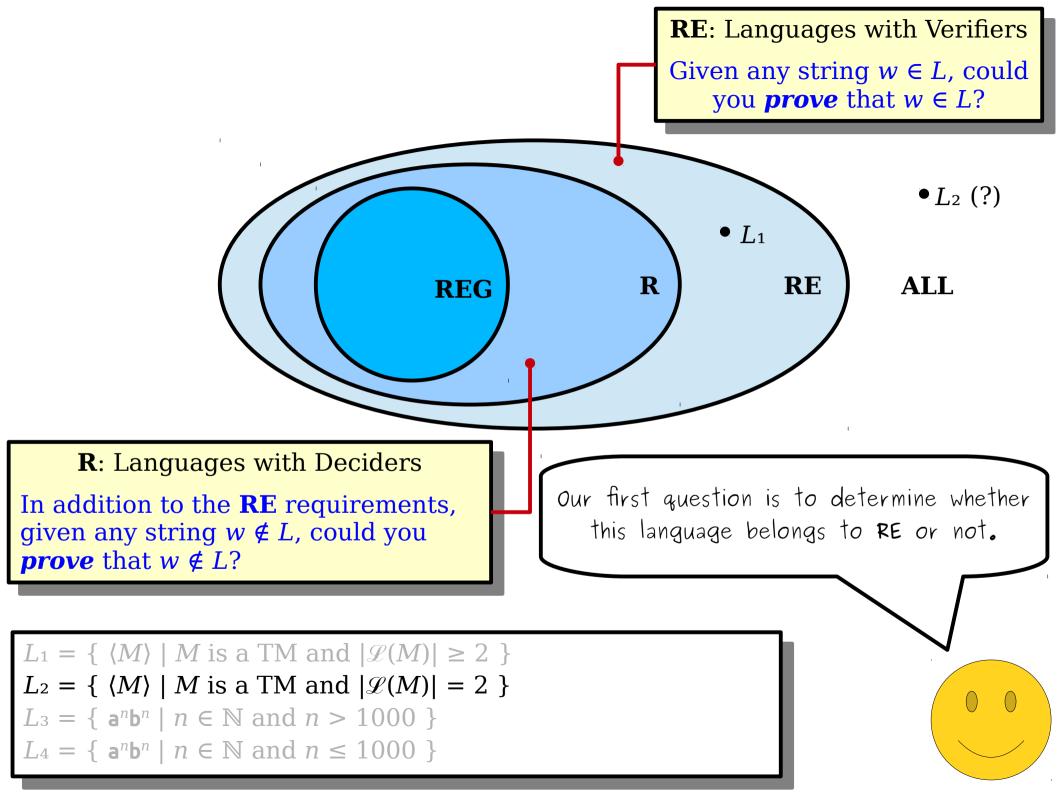
R: Languages with Deciders

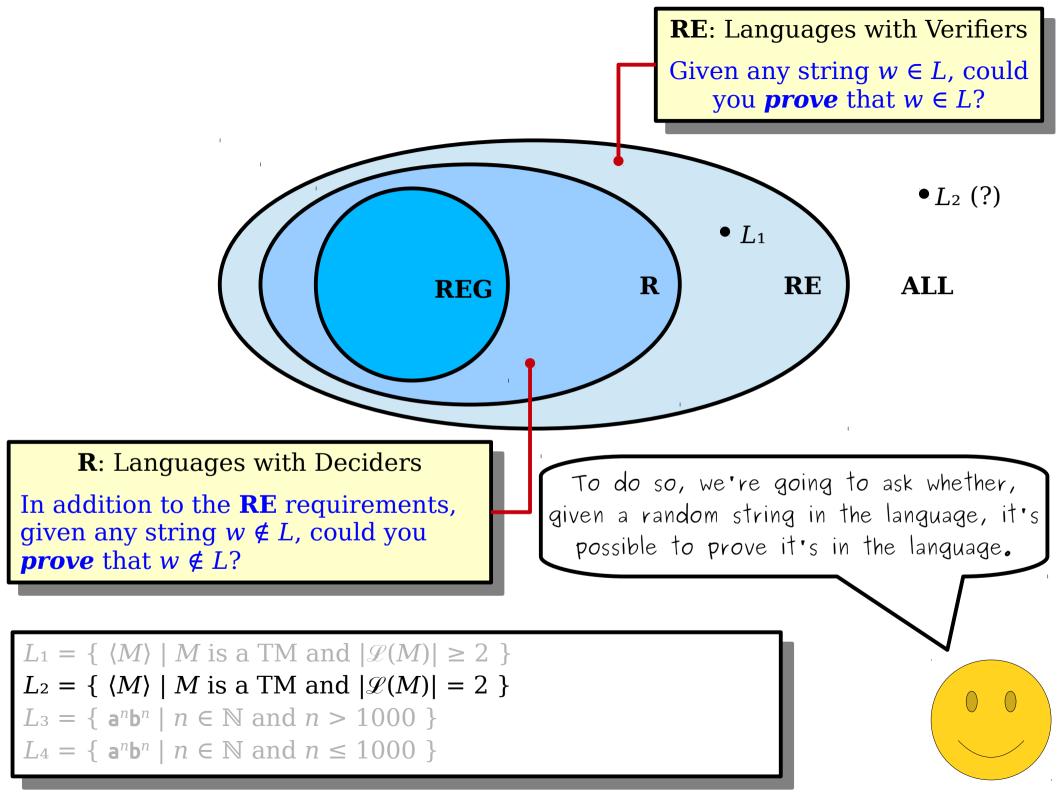
In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

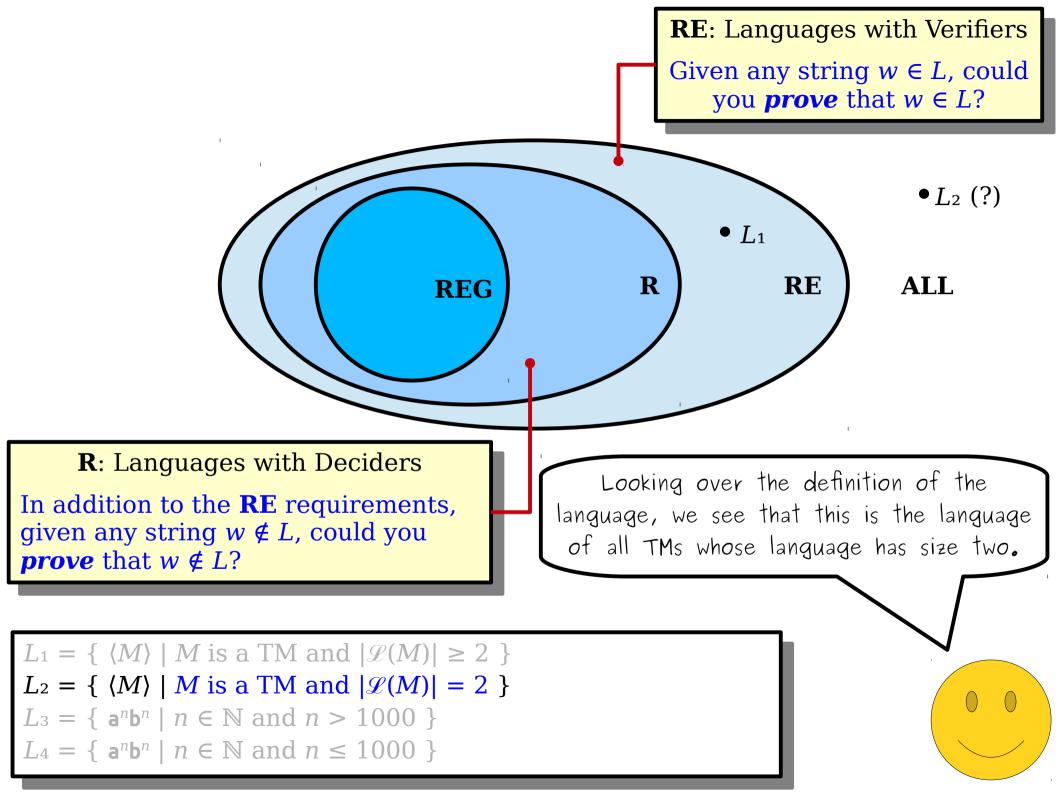
As before, we're going to start on the outside and move inward. Initially, we won't make any assumptions about where this particular language goes.

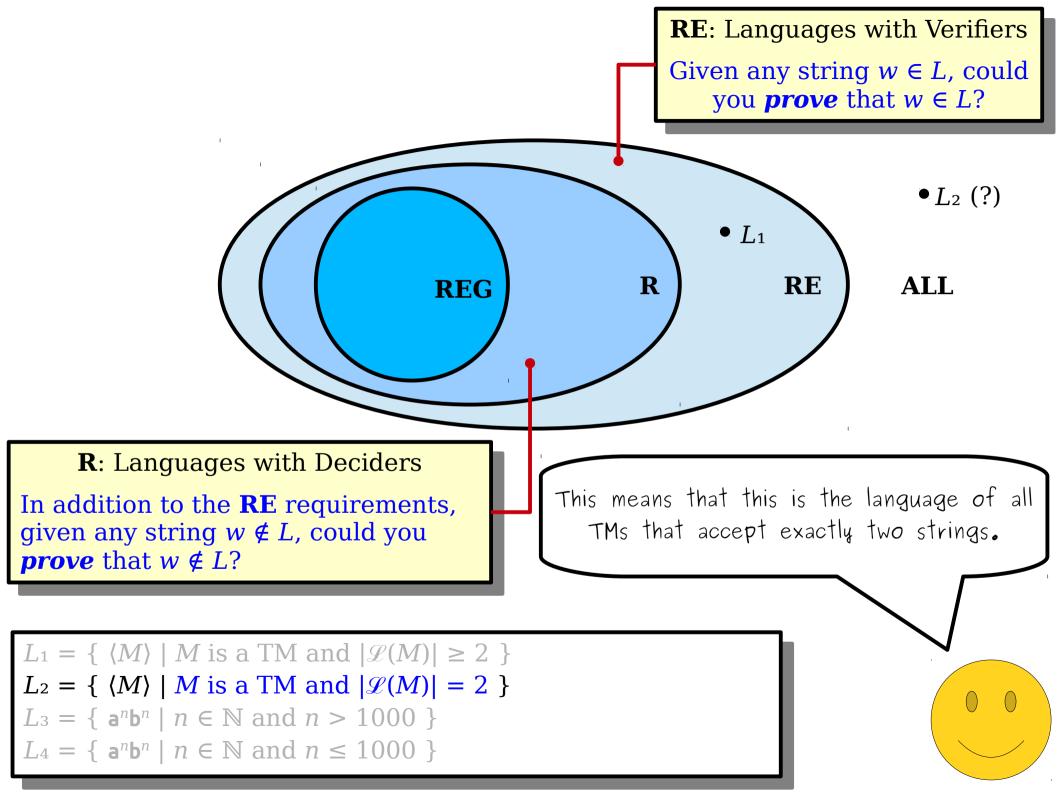
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

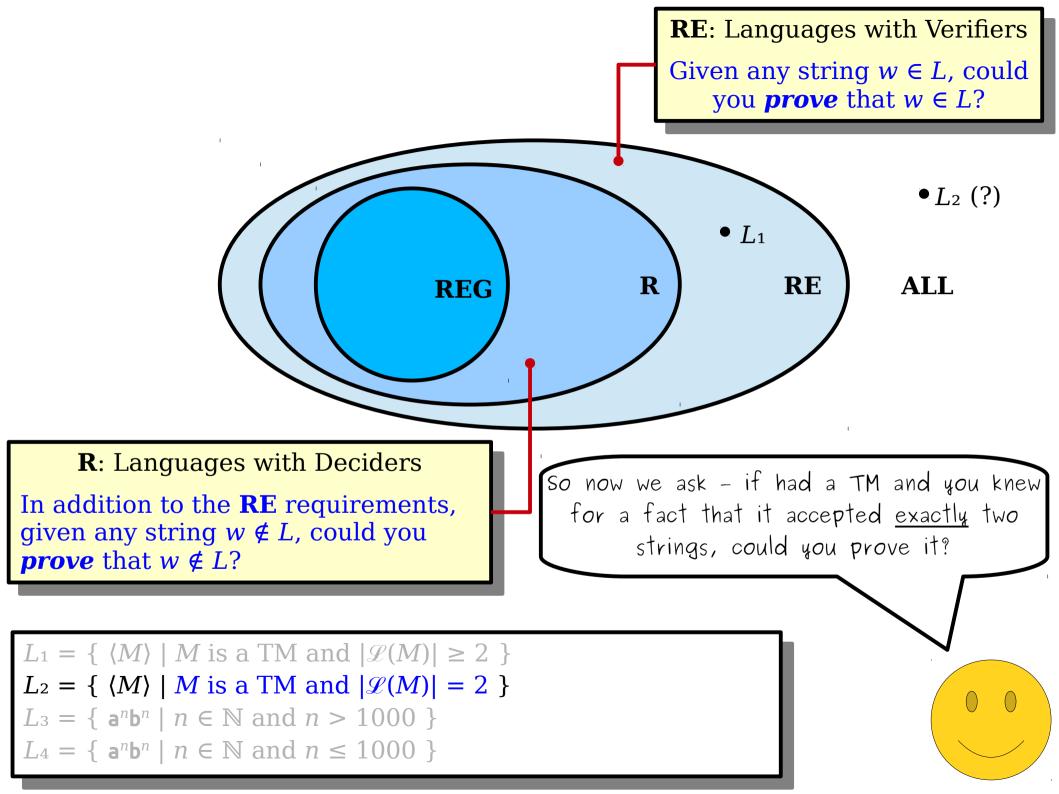


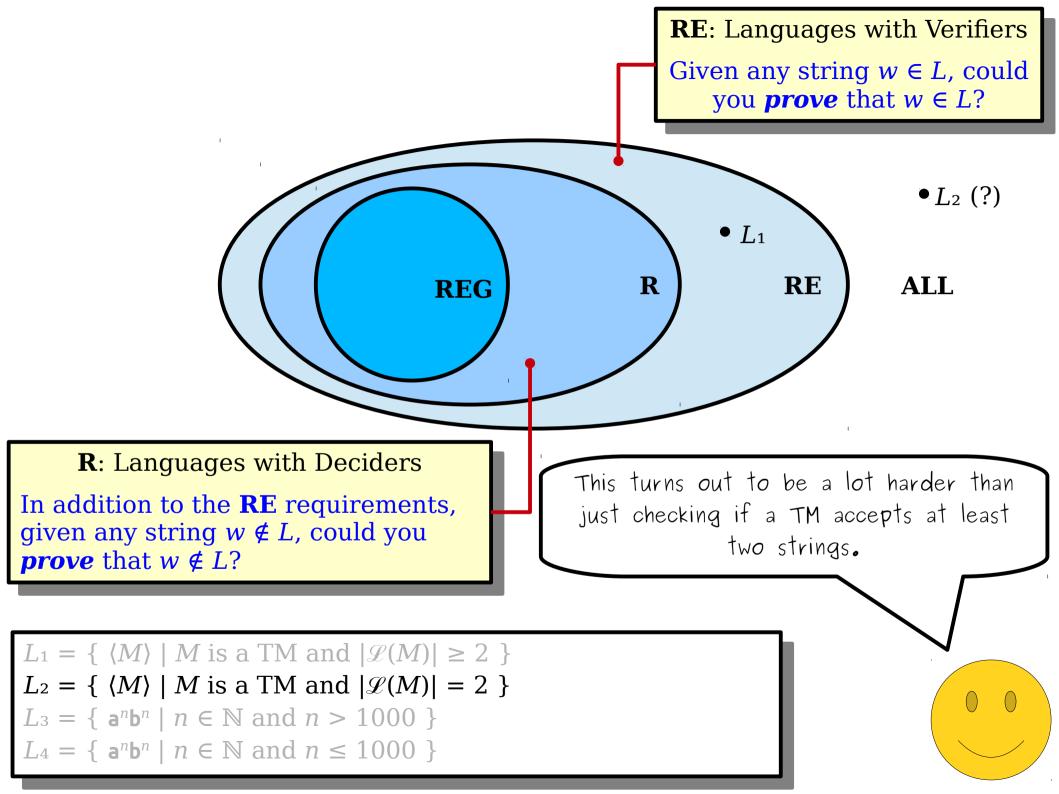


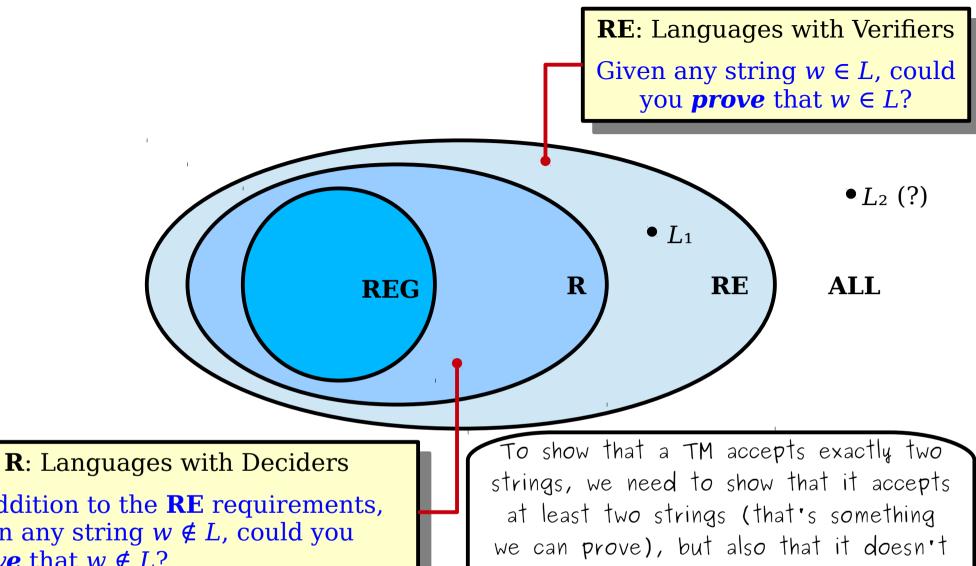










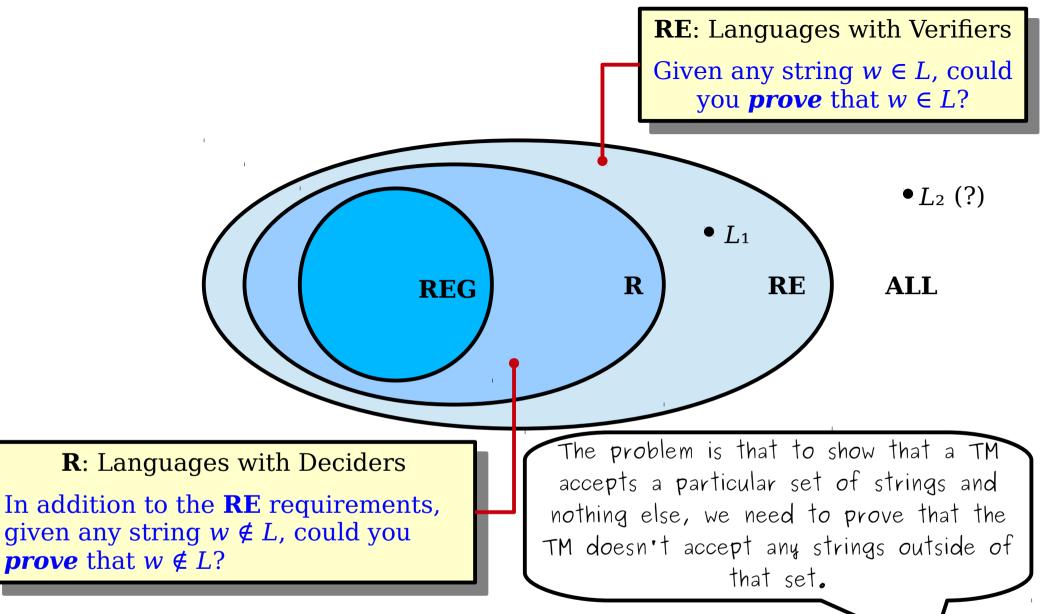


In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

accept anything else.

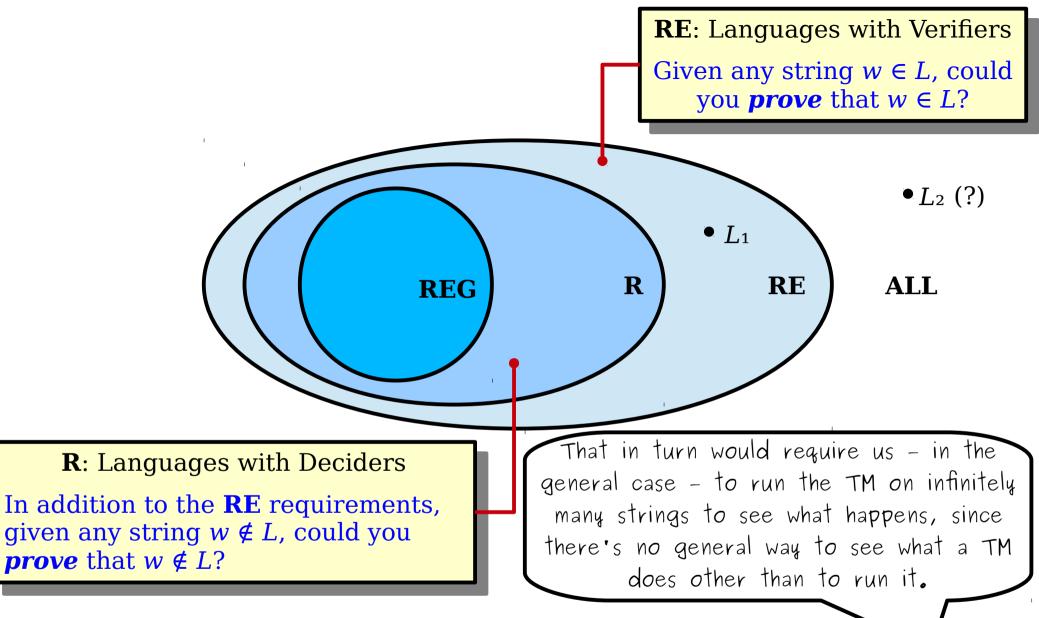
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ a^n b^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





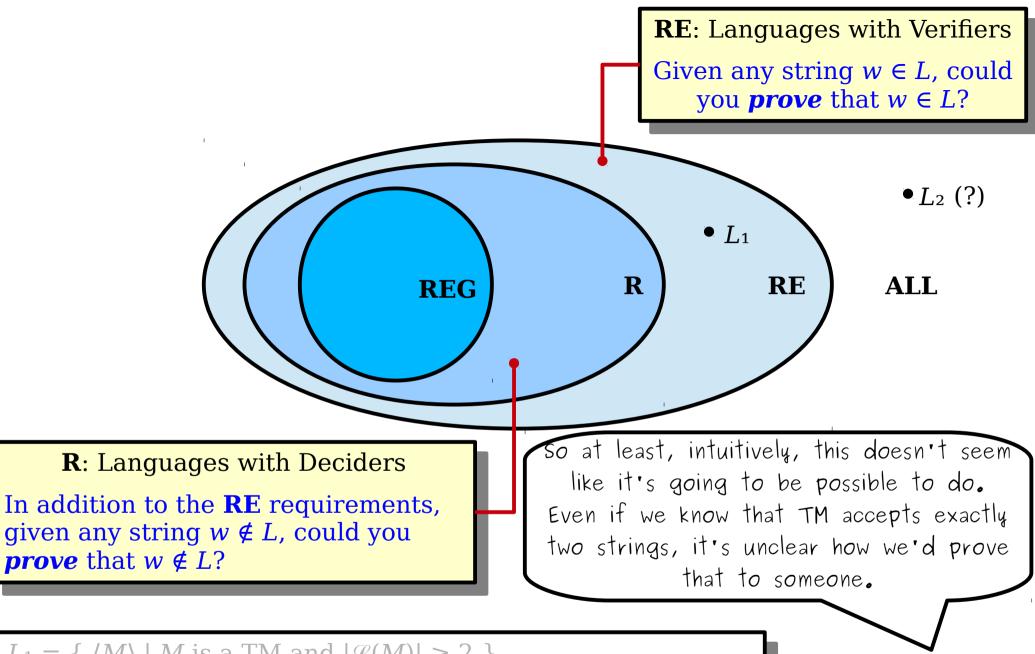
 $L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}$ $L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}$ $L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}$ $L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}$





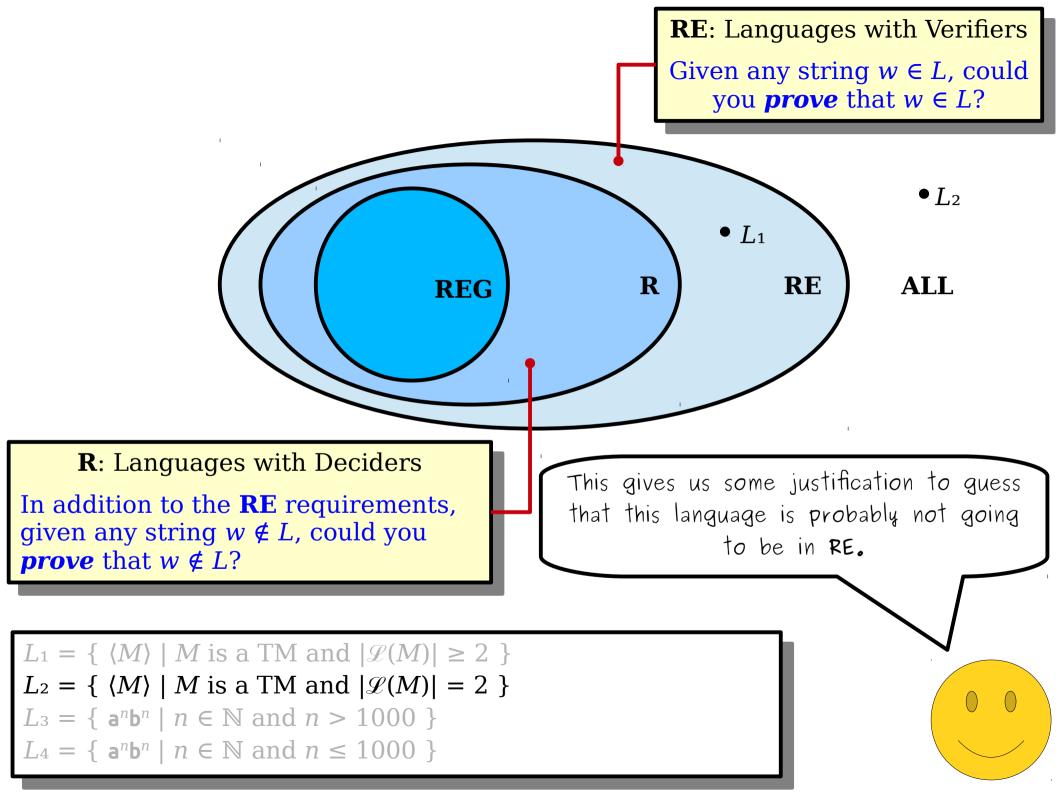
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

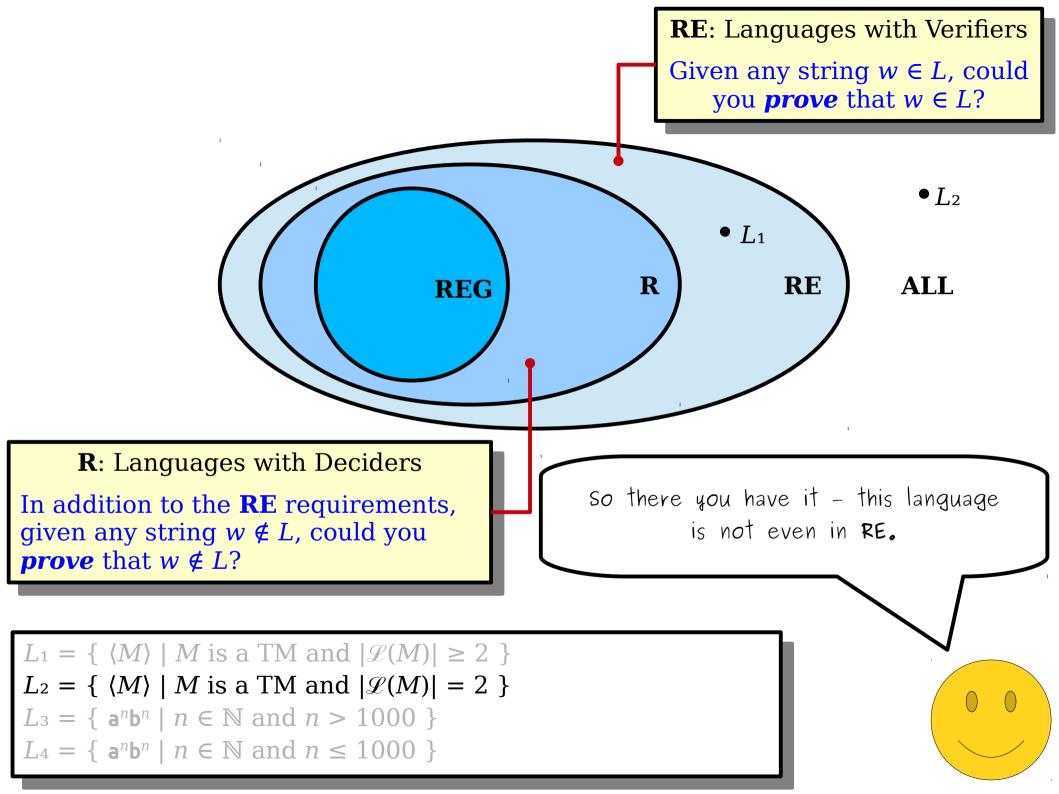


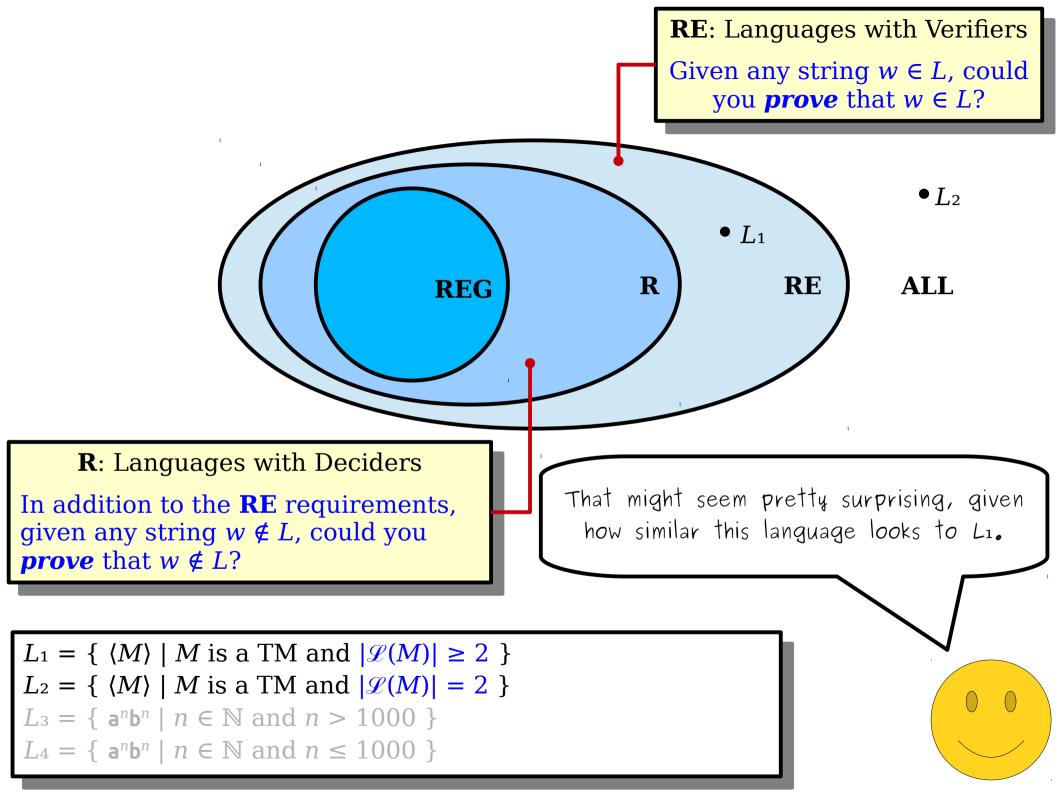


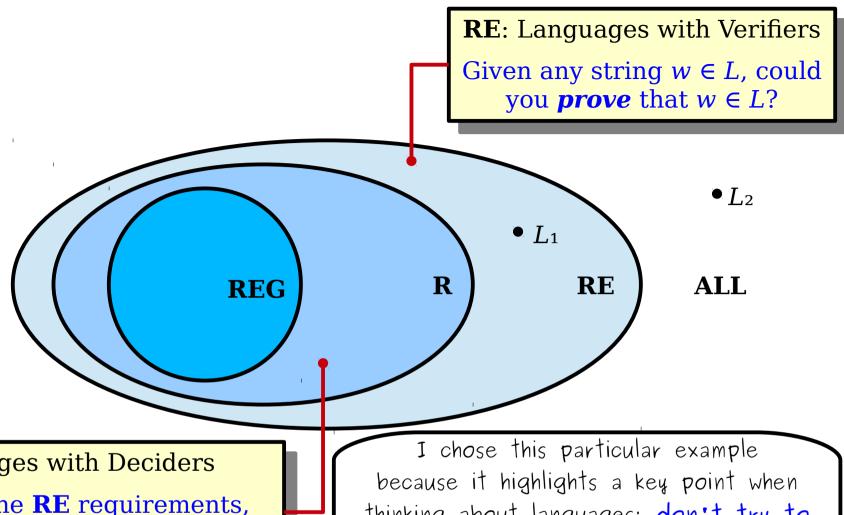
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```









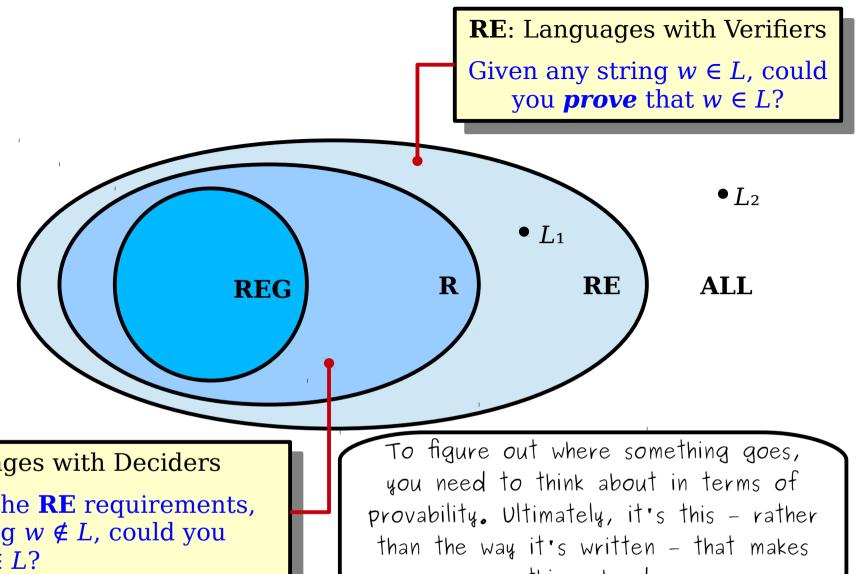


In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

thinking about languages: don't try to place a language in the diagram just based on its description.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}
L_3 = \{ a^n b^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```



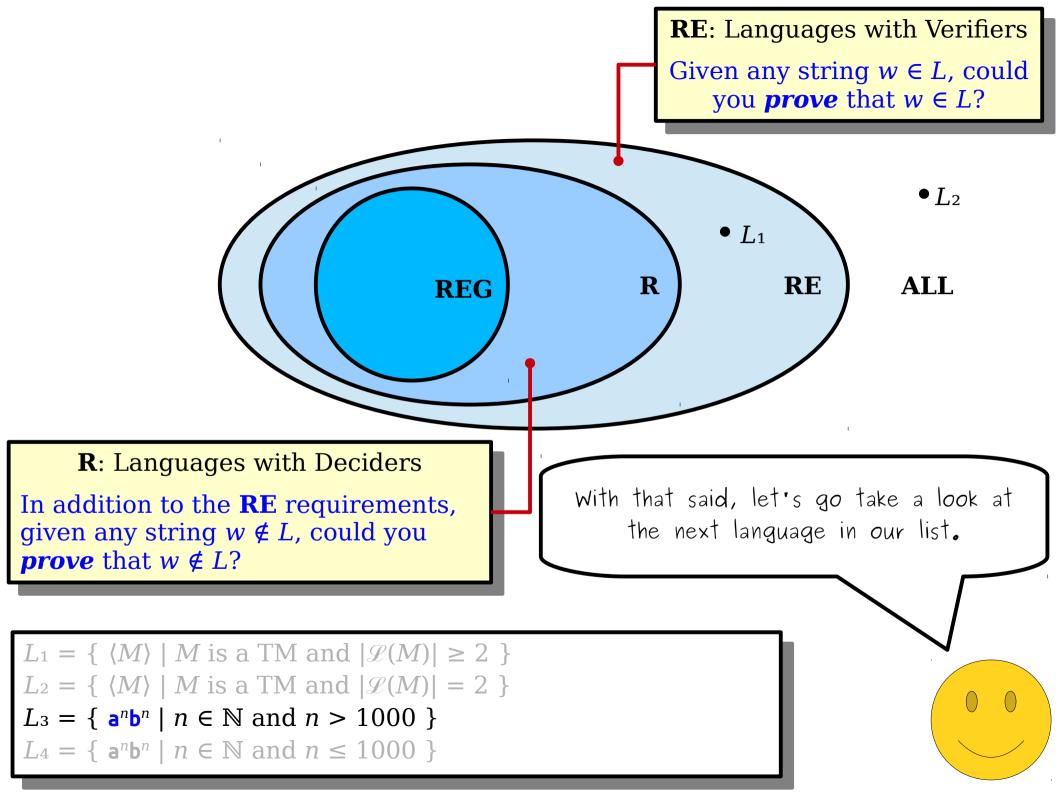


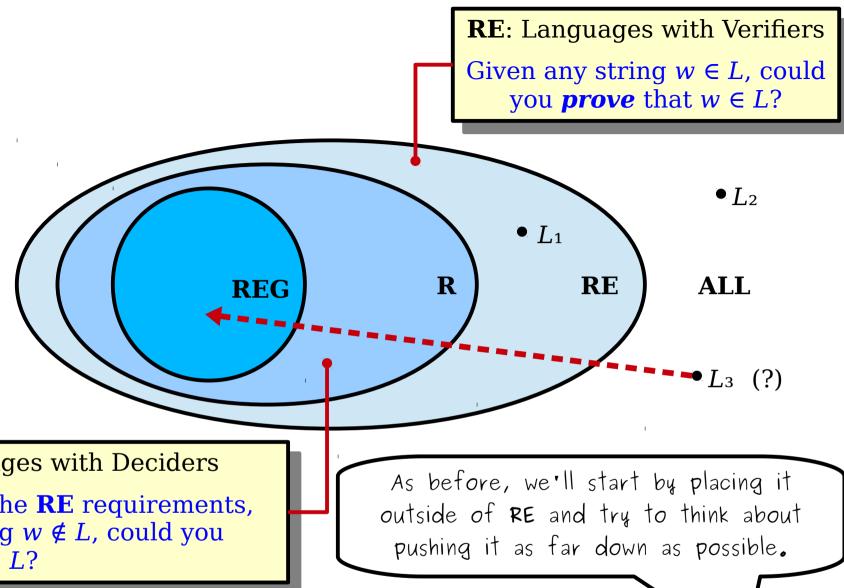
In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

things hard.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}
L_3 = \{ a^n b^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```



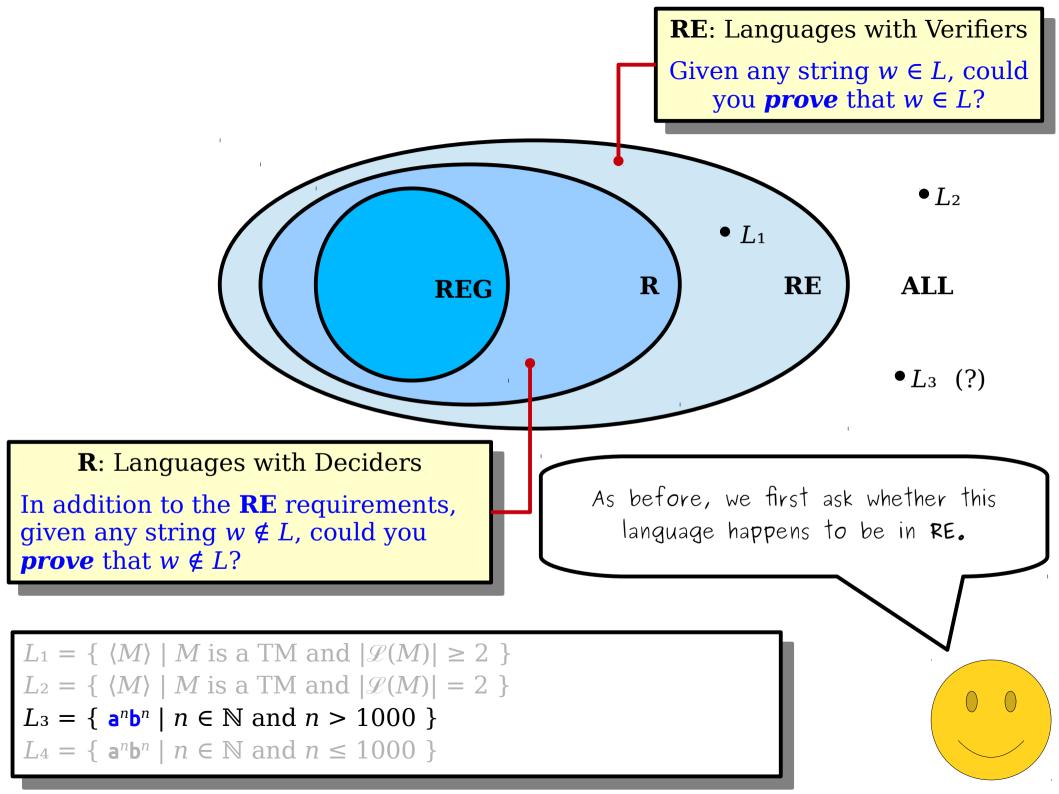


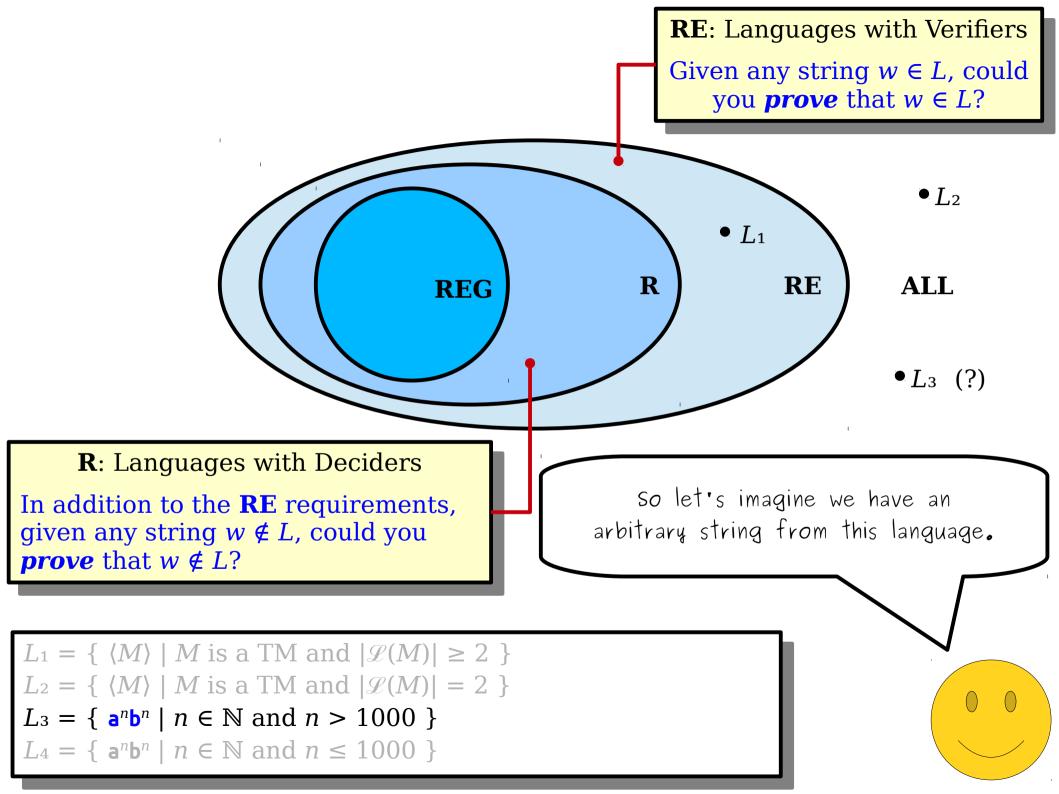


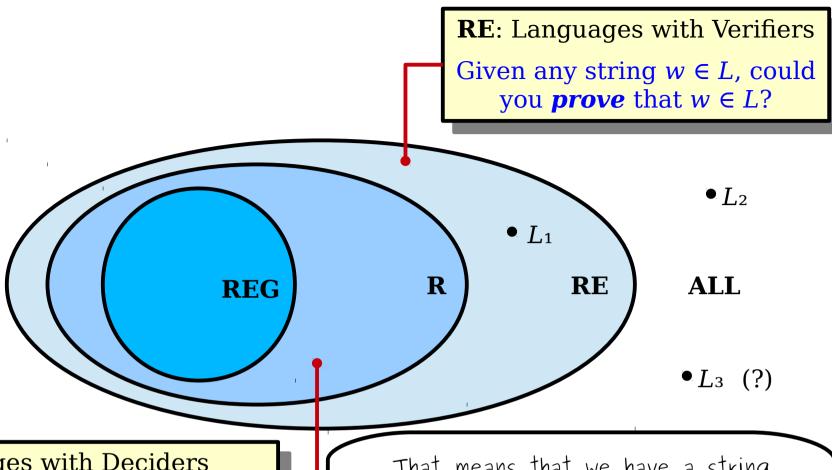
In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ a^n b^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```







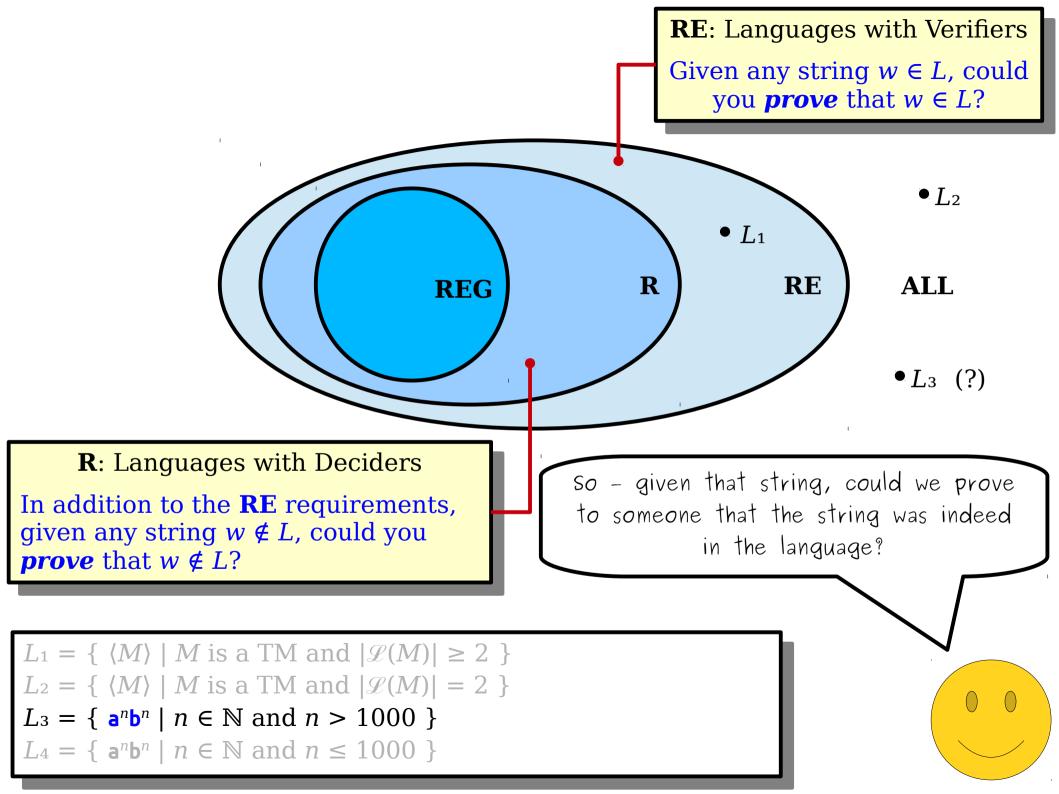


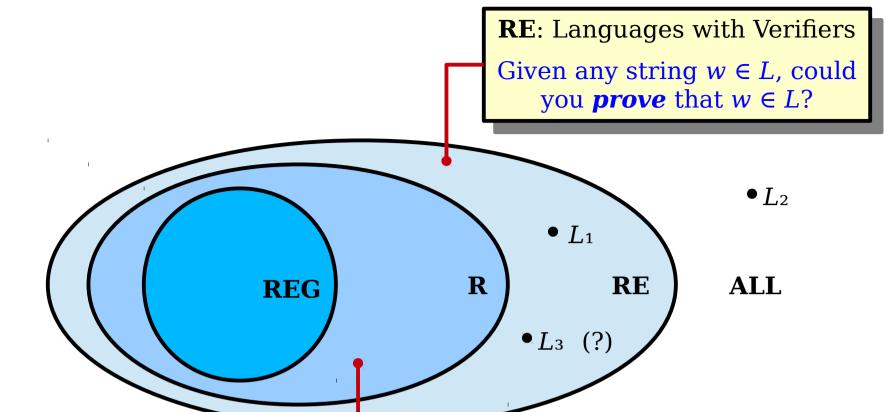
In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

That means that we have a string of the form anb with at least 2,002 characters in it (at least 1,001 a's and at least 1,001 b's.)

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





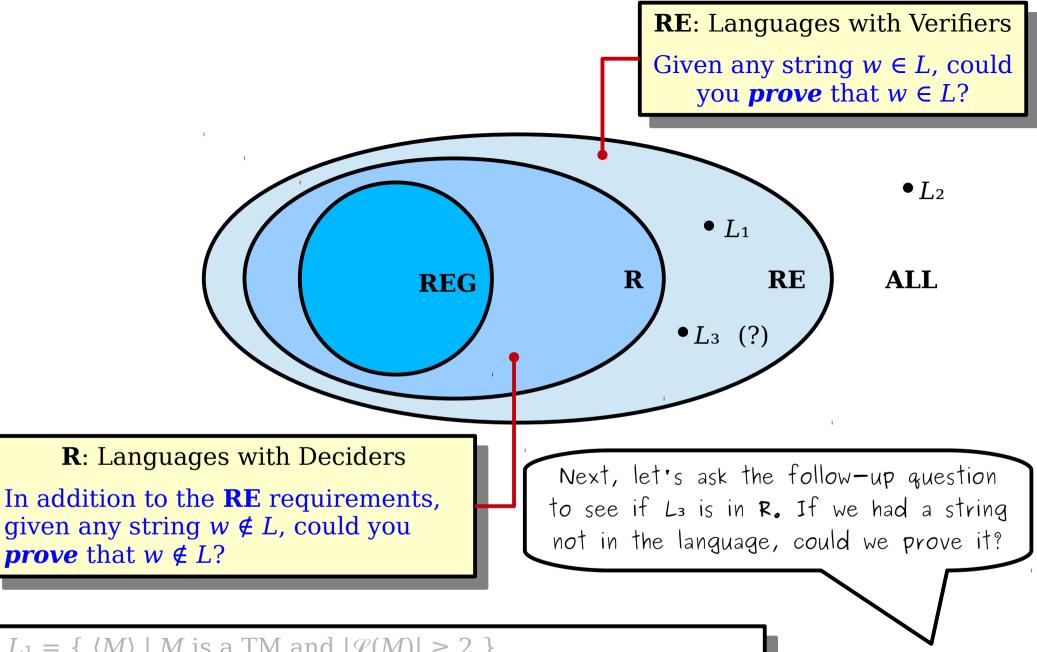


In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

Sure! We could just count up the a's, count up the b's, show that there are the same number, and show that there's at least 1,000.

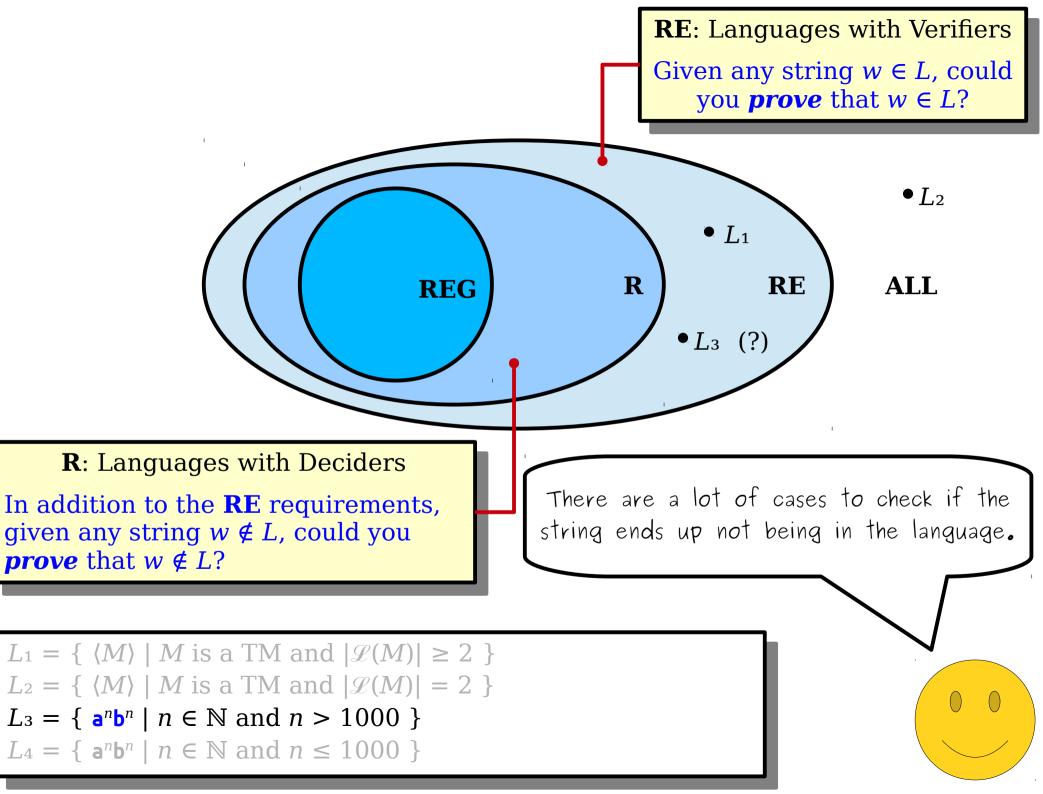
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

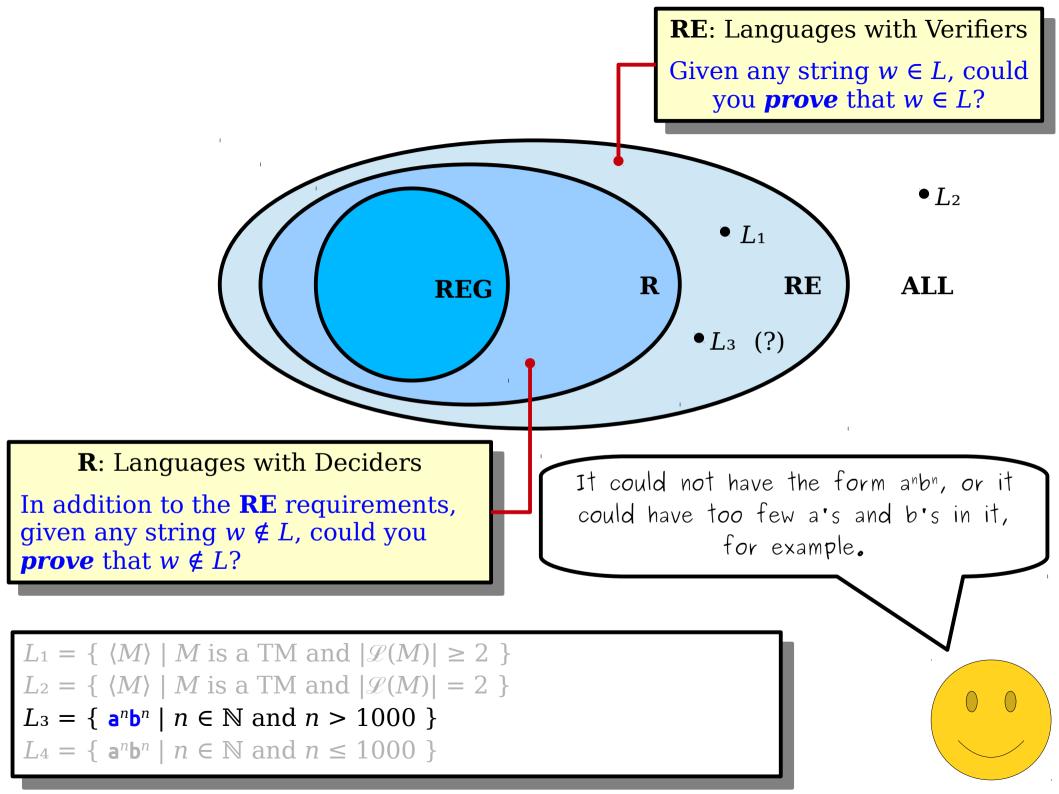


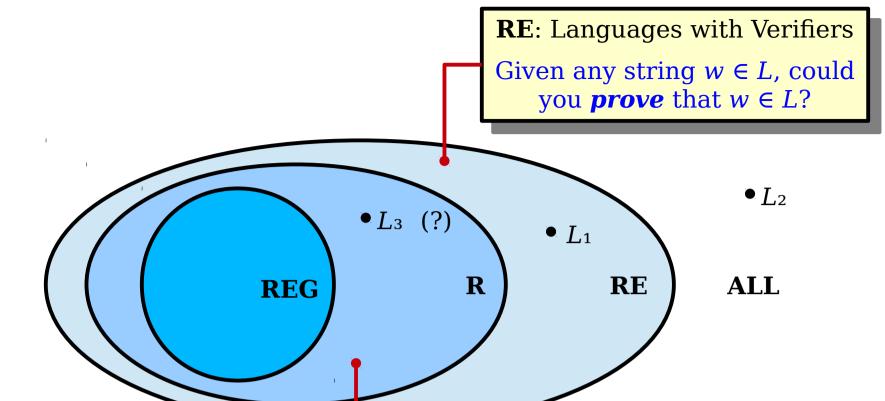


```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```







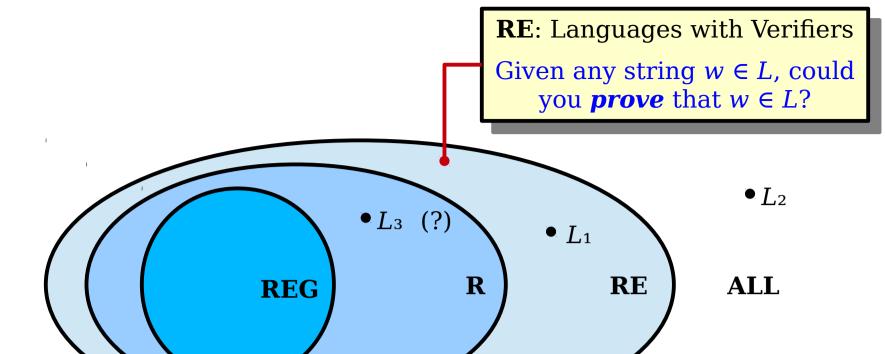


In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

However, all of those cases are really easy to check. We either show that it has the wrong form or show that it doesn't have enough characters in it.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

Okay, things are looking good here! We know that this language is decidable. As our final step, we need to ask whether or not it's regular.

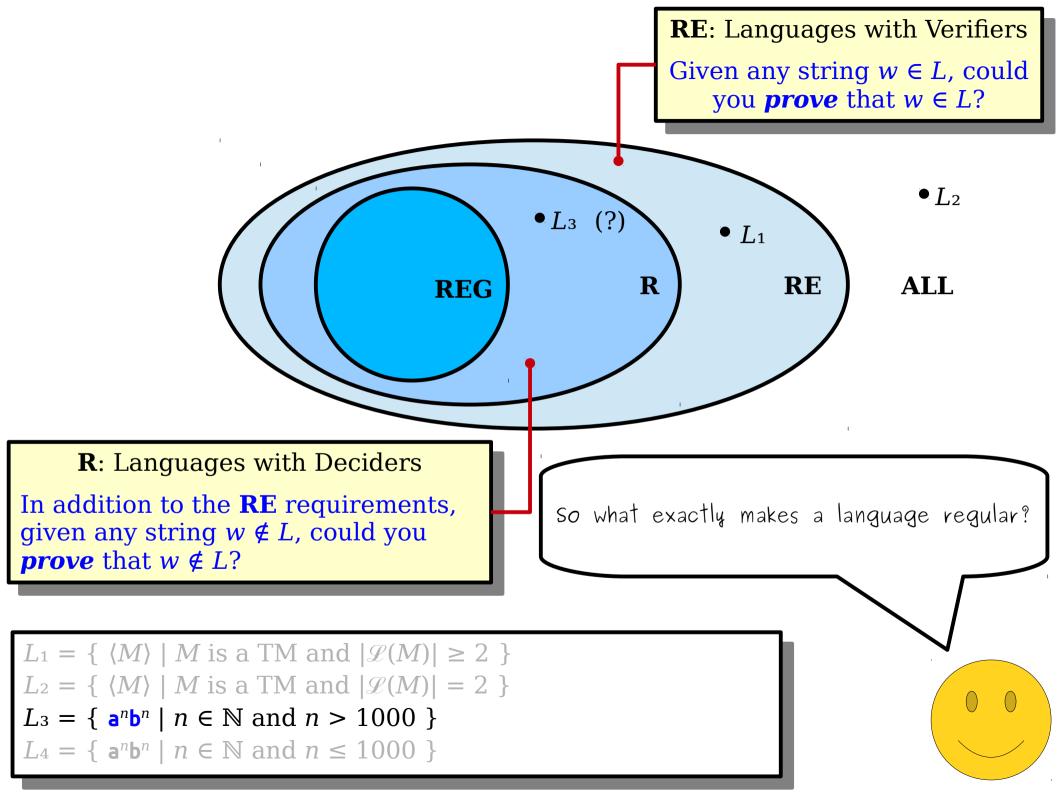
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}

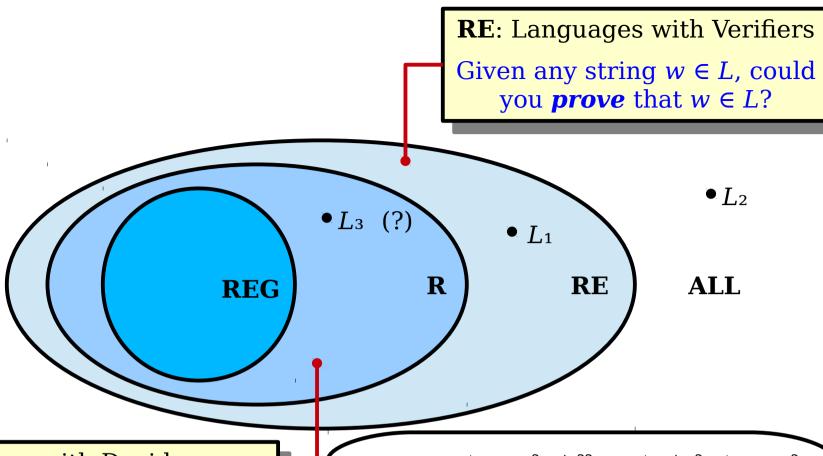
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}

L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}

L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```





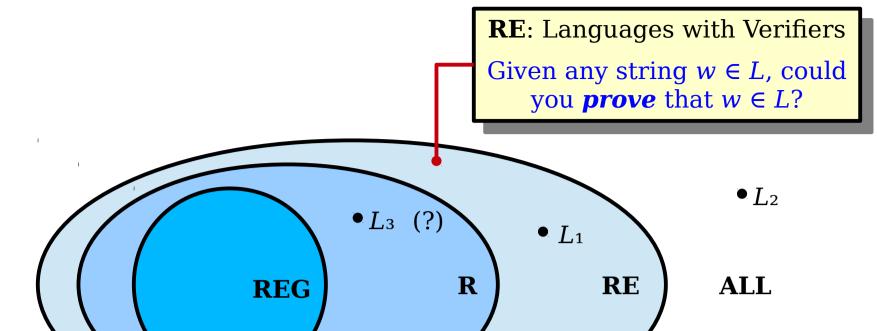


In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

We have a ton of different definitions for regular languages - they're the languages of DFAs, NFAs, regexes, and right-linear grammars.

```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathscr{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```



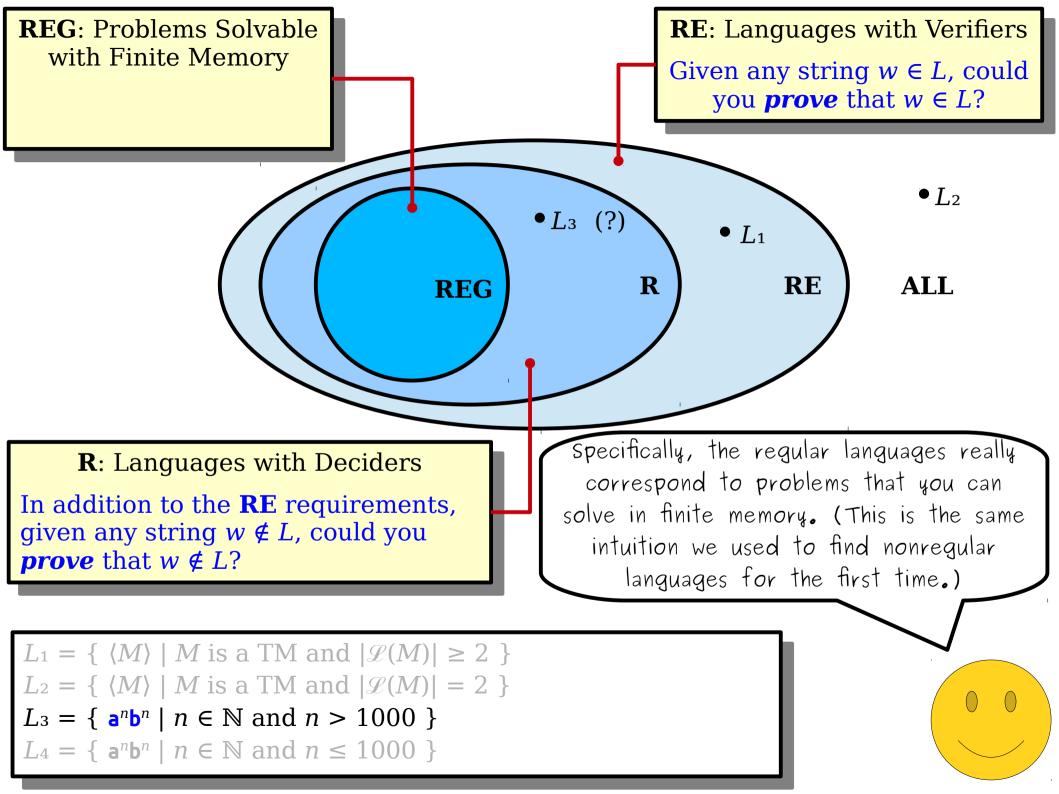


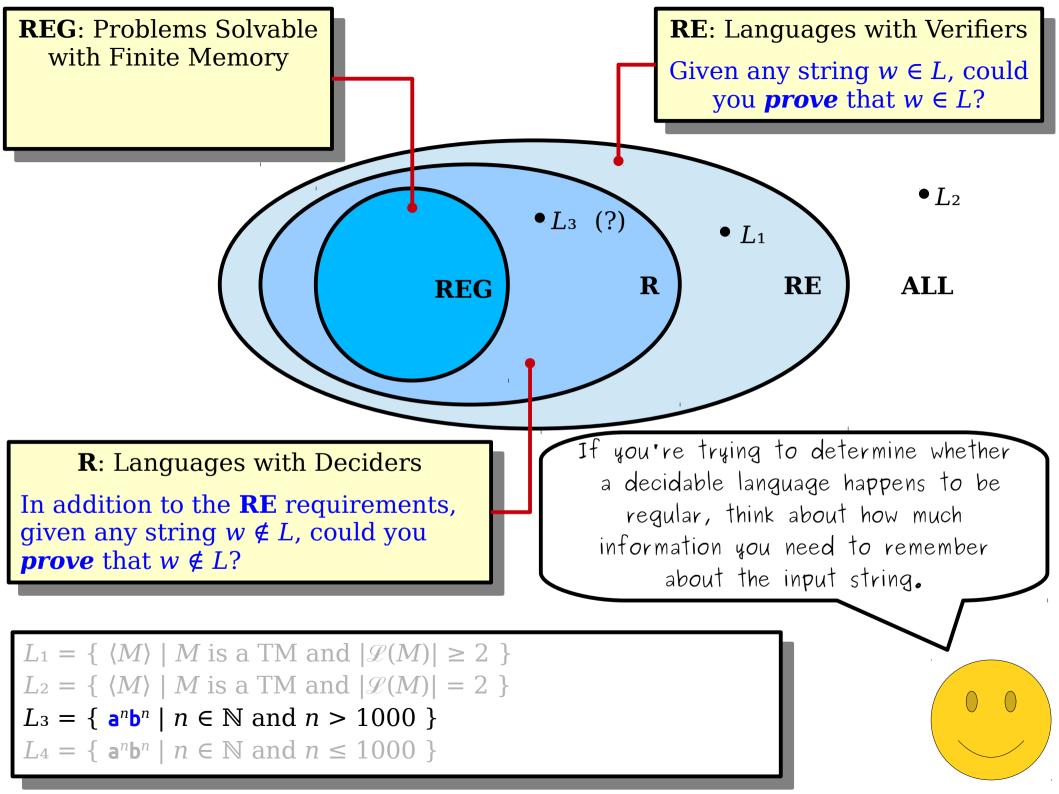
In addition to the **RE** requirements, given any string $w \notin L$, could you **prove** that $w \notin L$?

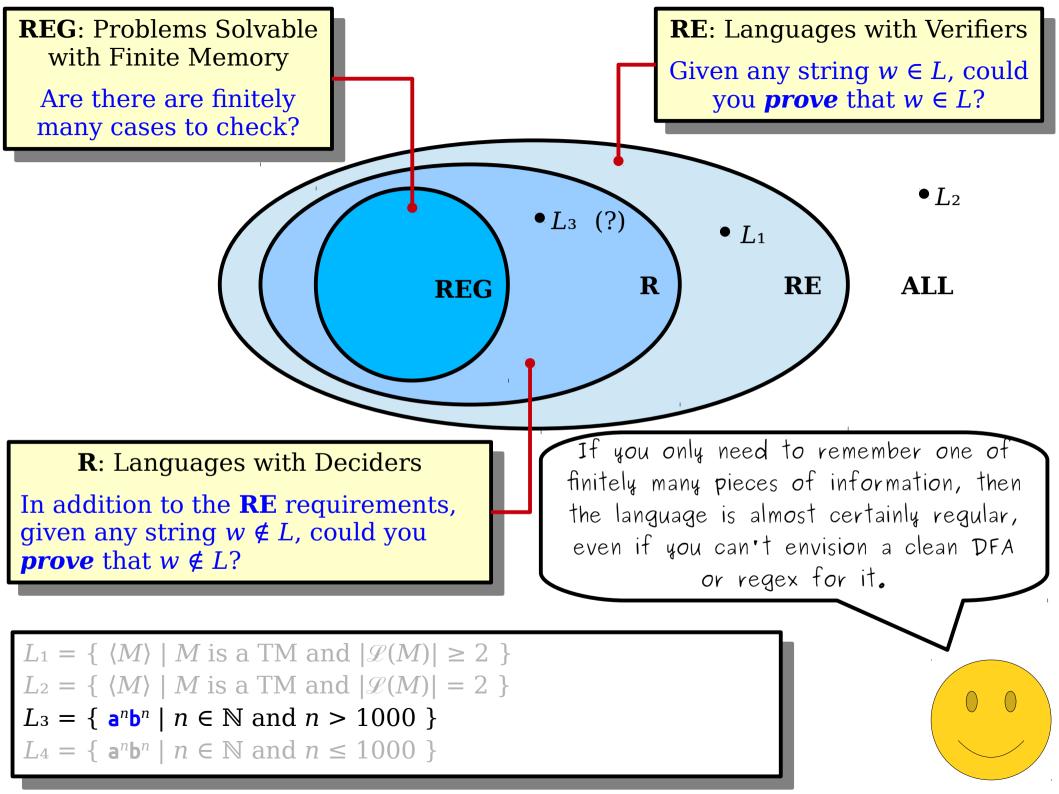
But, as with R and RE, I think there's a much better intuition to have about the regular languages that makes it easier to see whether something is regular.

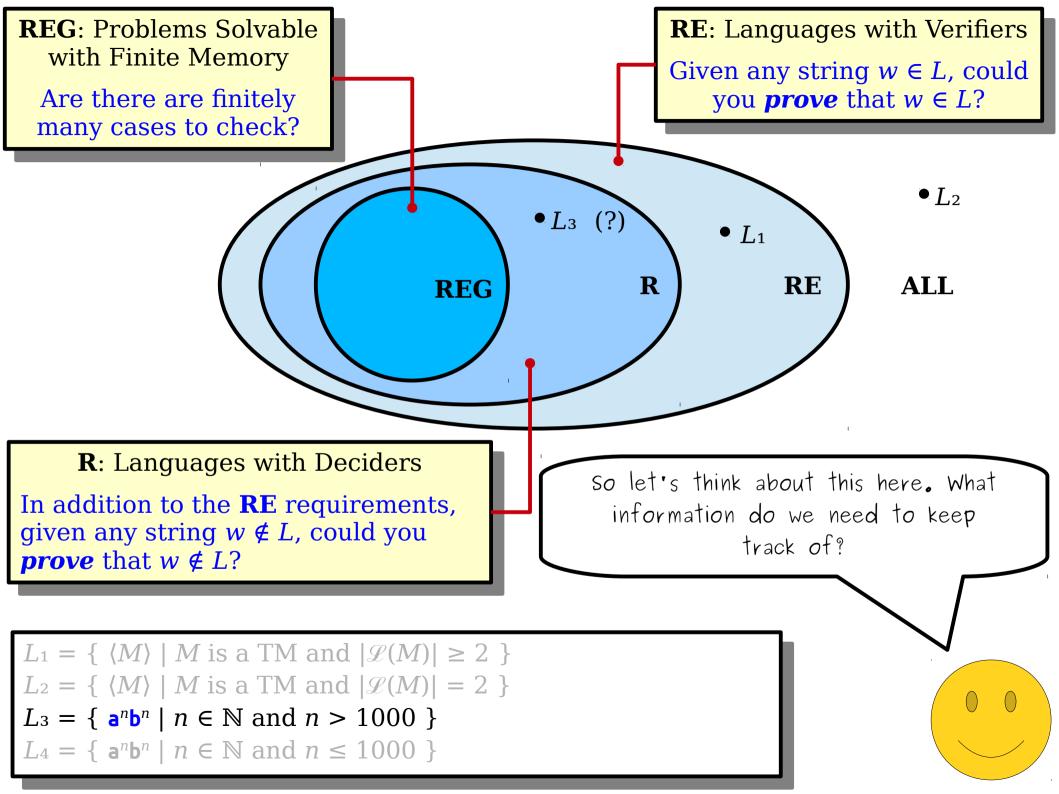
```
L_1 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | \geq 2 \}
L_2 = \{ \langle M \rangle \mid M \text{ is a TM and } | \mathcal{L}(M) | = 2 \}
L_3 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n > 1000 \}
L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}
```

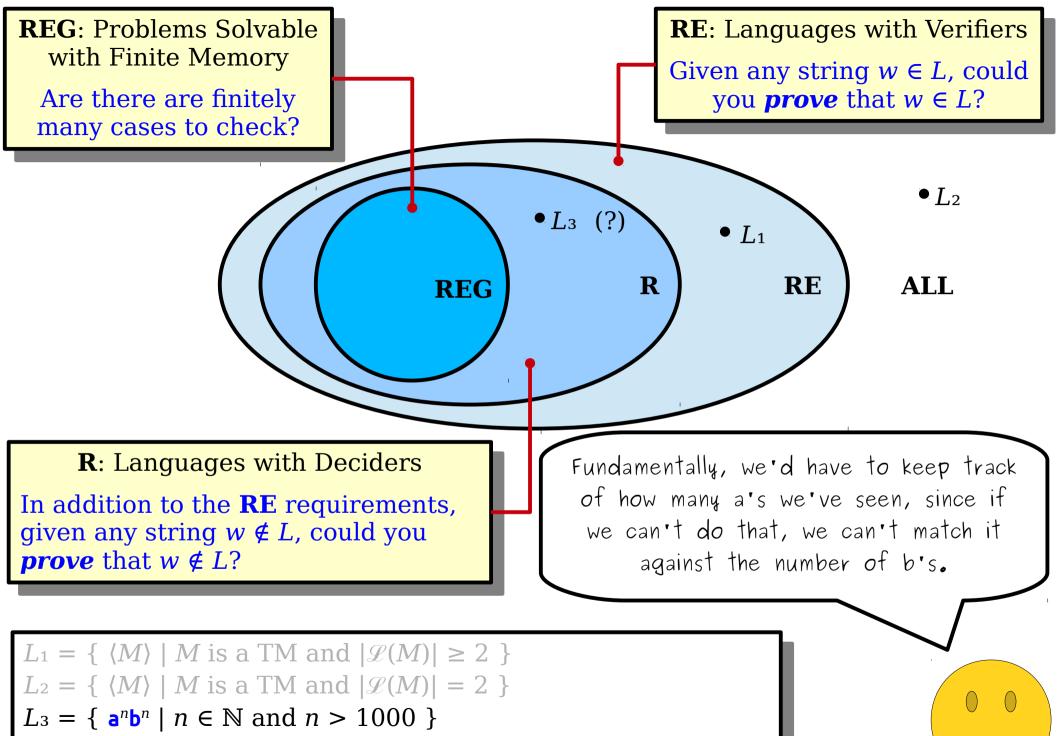




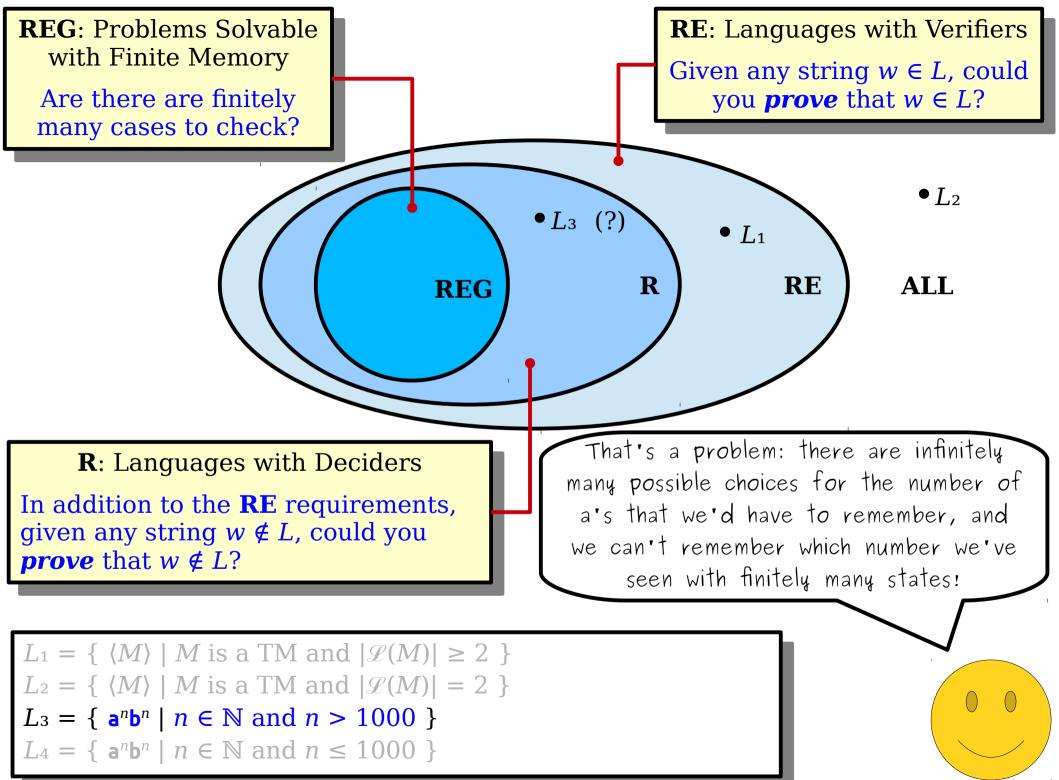


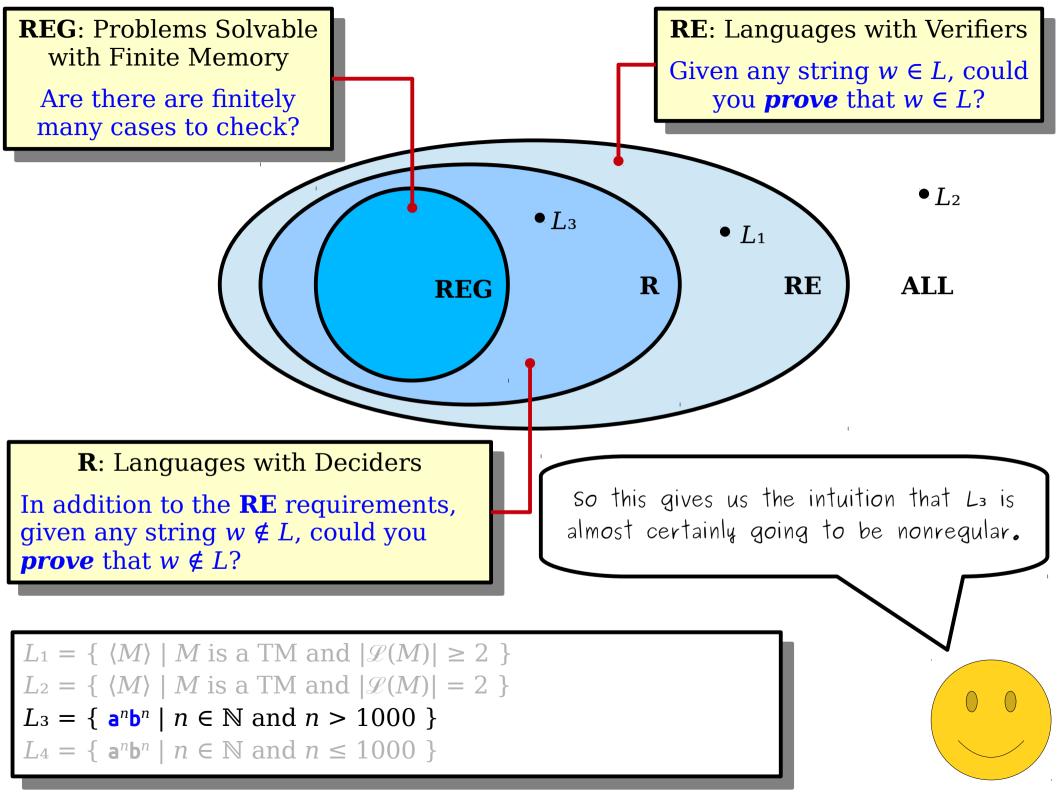


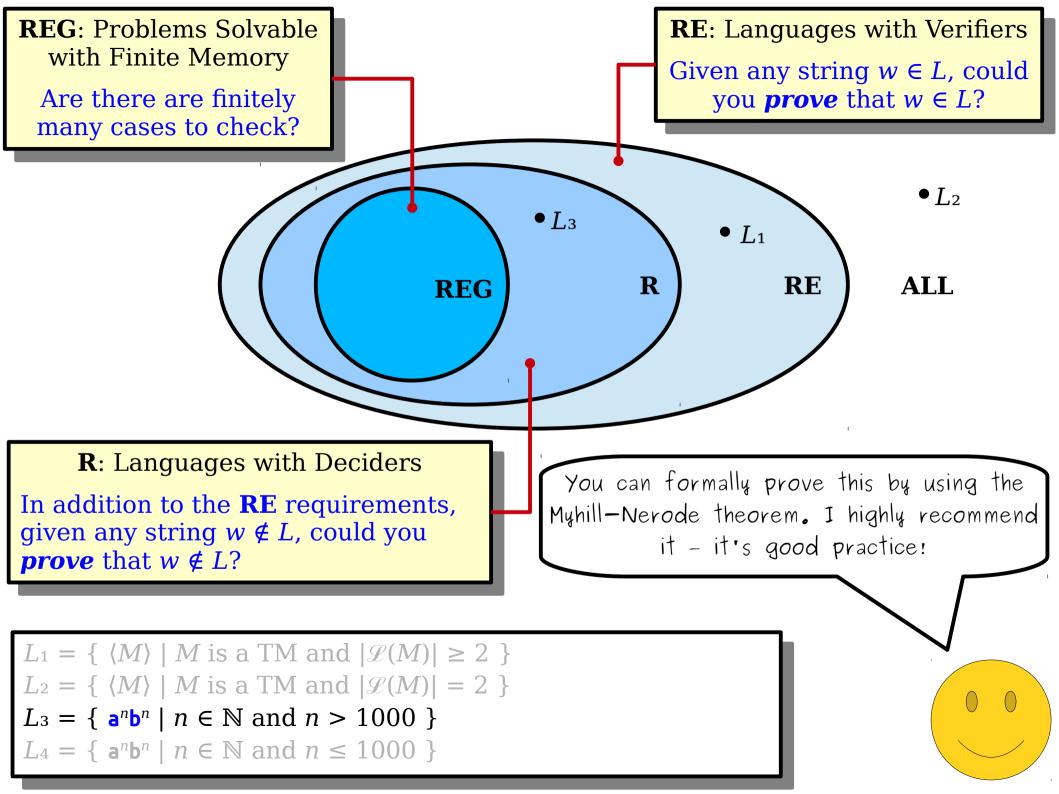


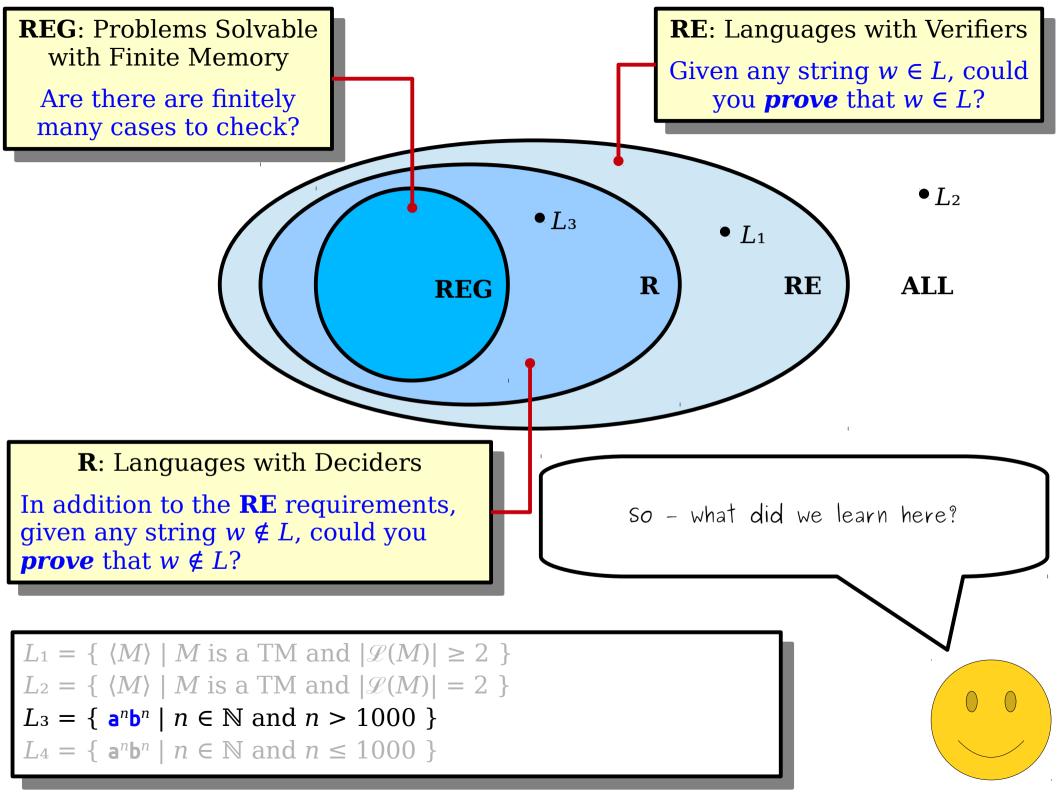


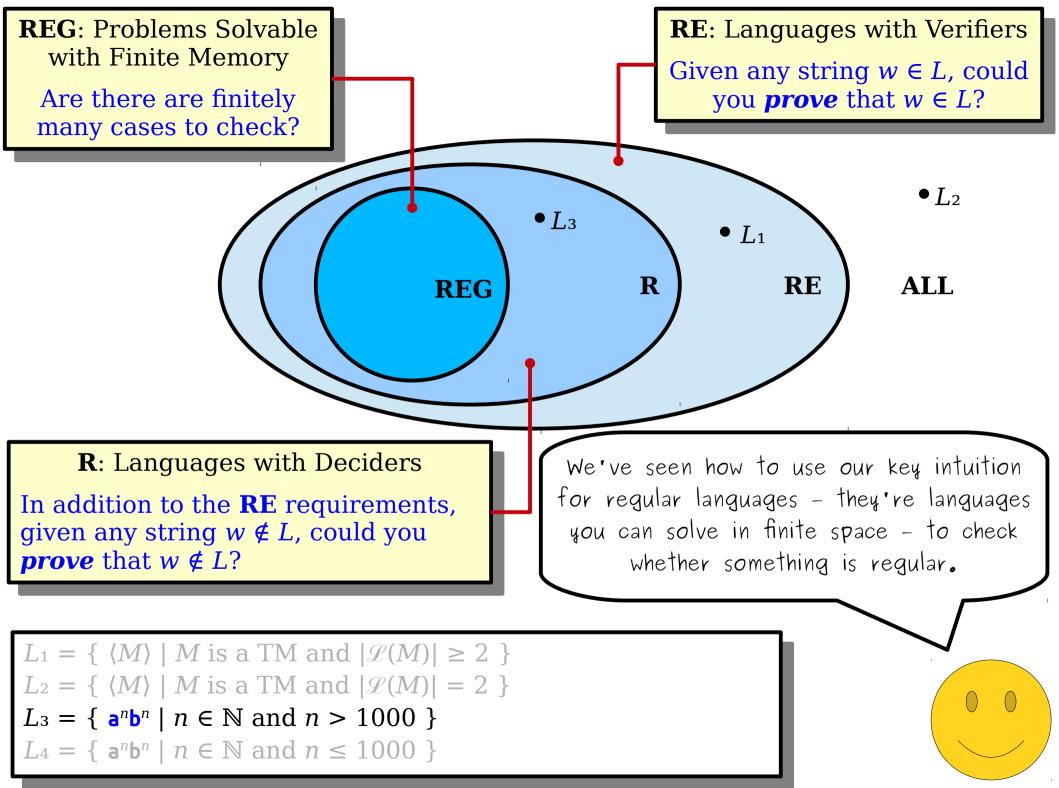
 $L_4 = \{ \mathbf{a}^n \mathbf{b}^n \mid n \in \mathbb{N} \text{ and } n \leq 1000 \}$

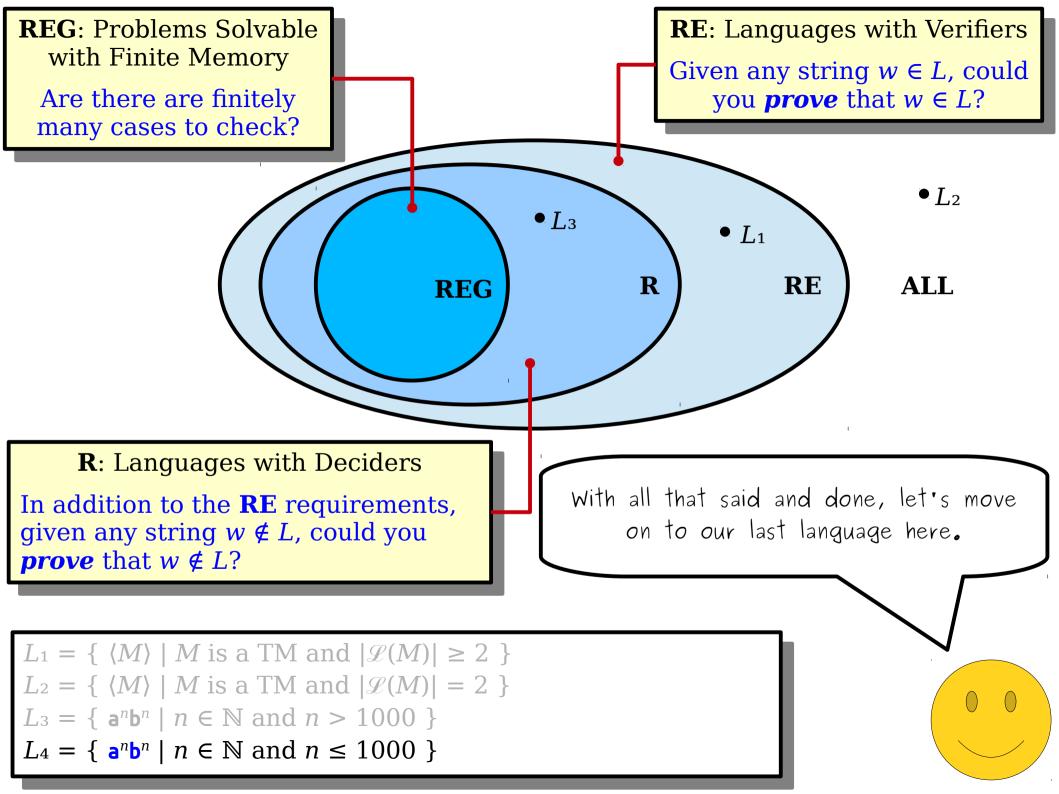


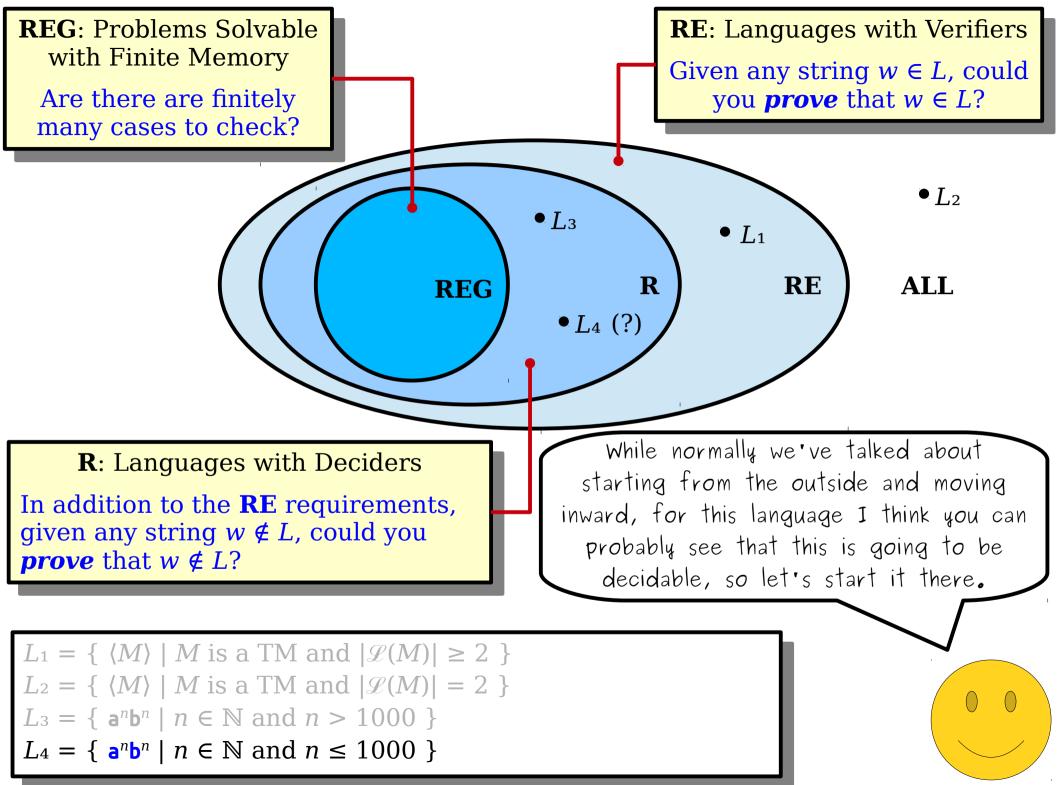


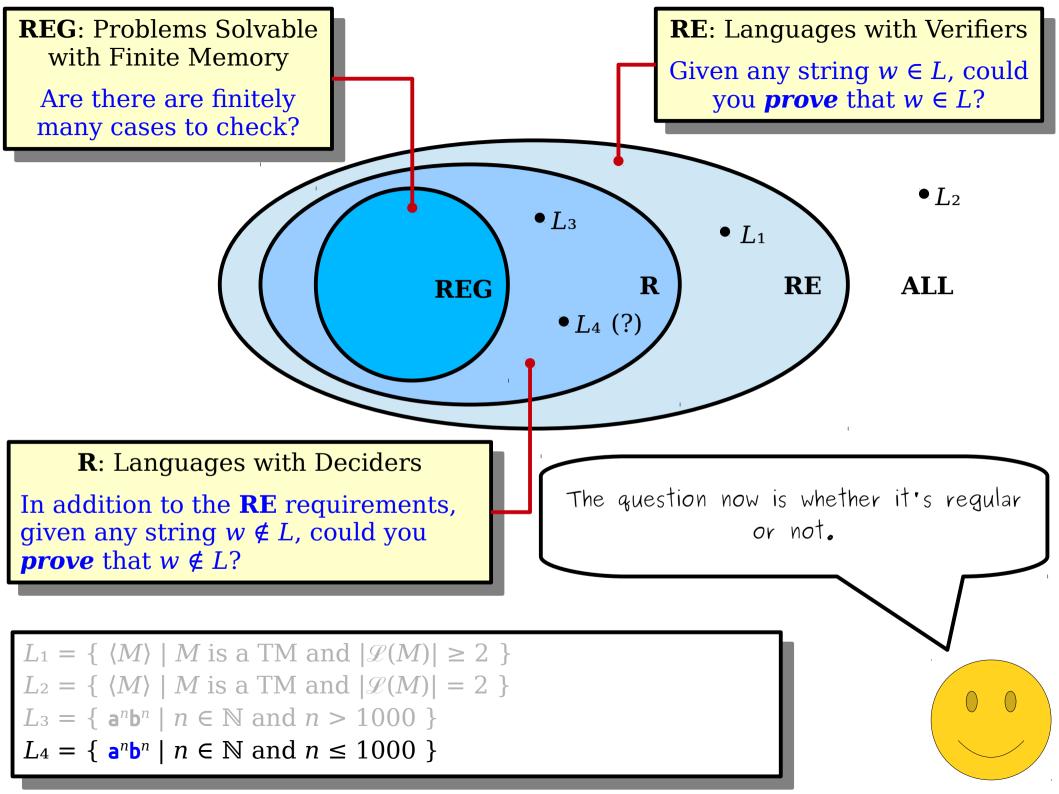


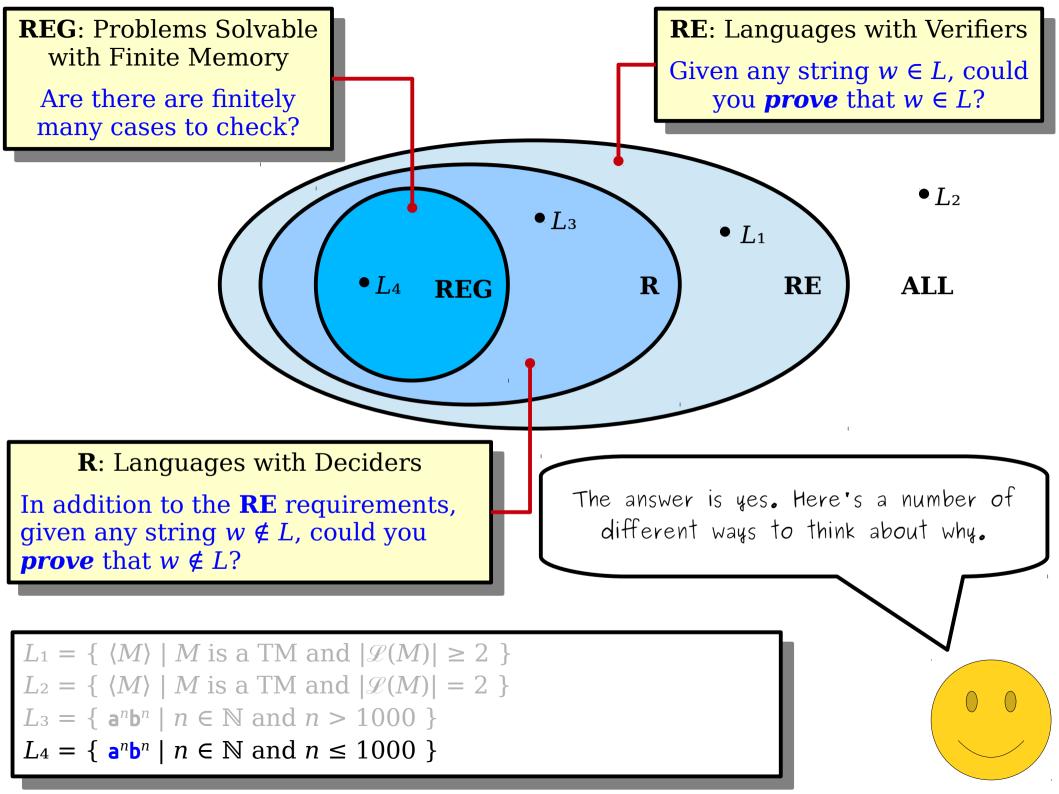


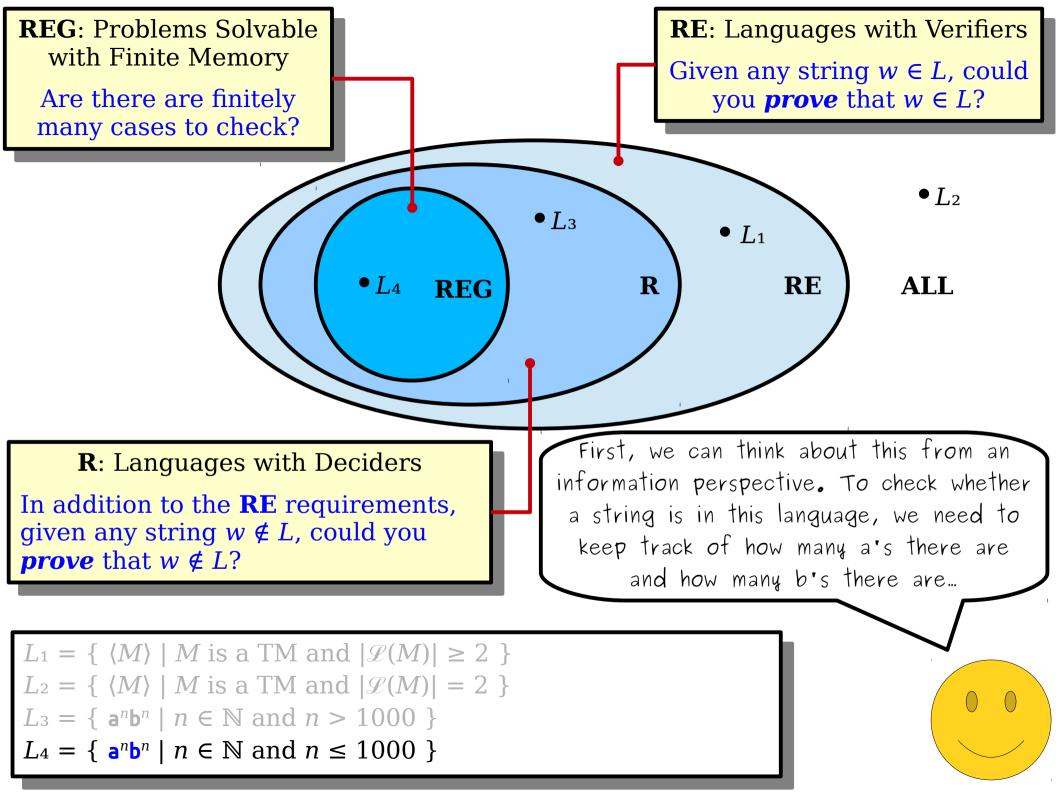


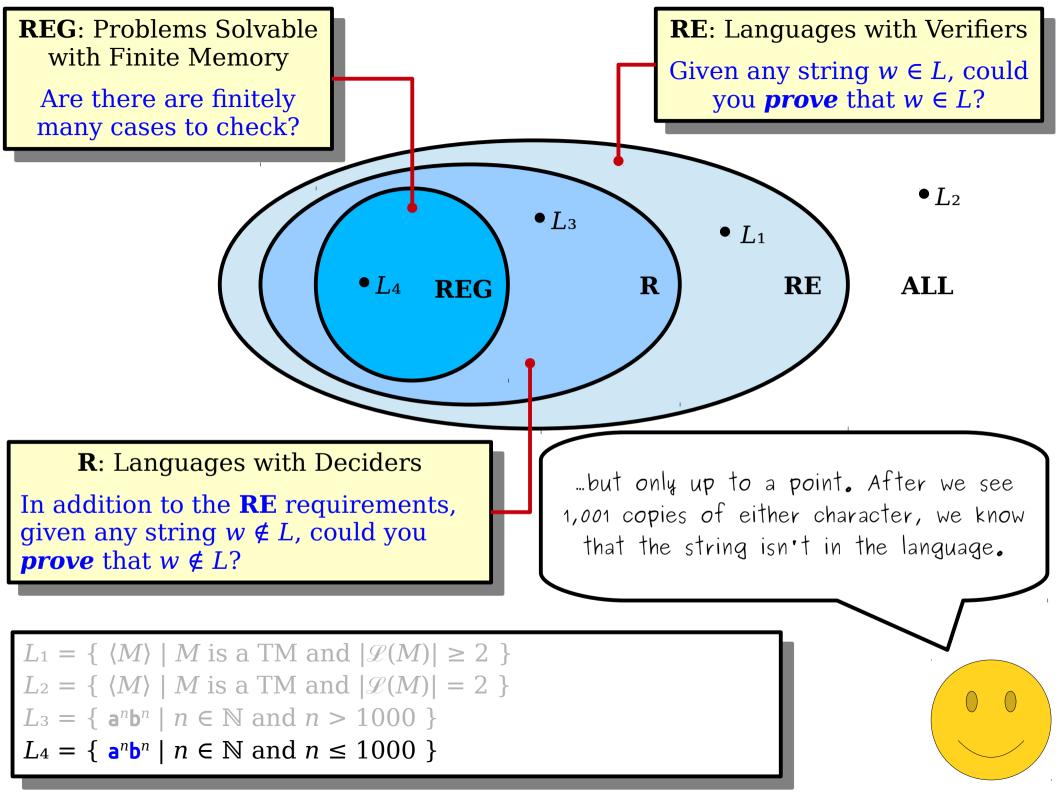


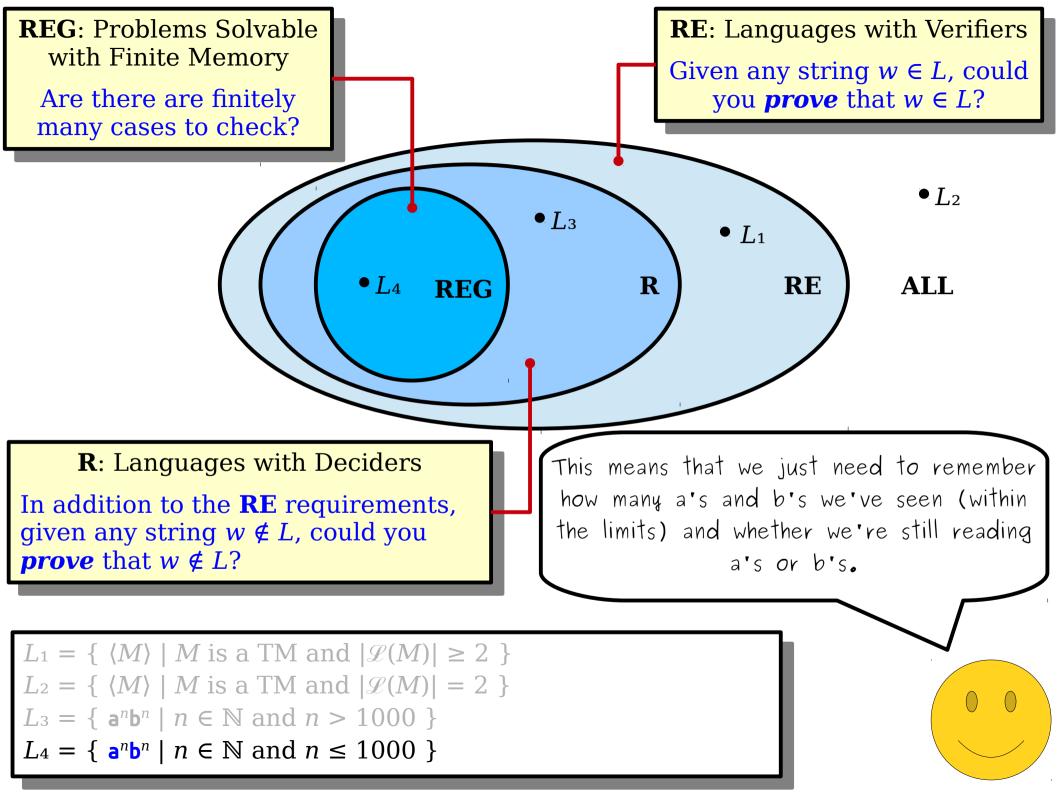


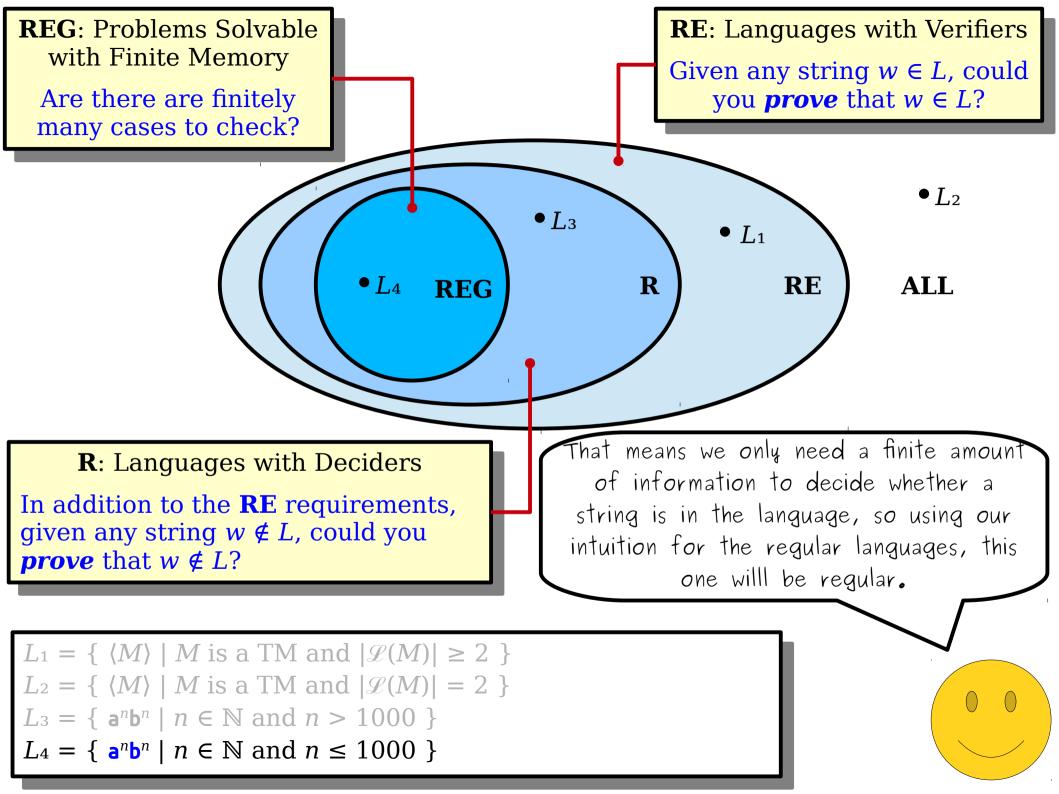


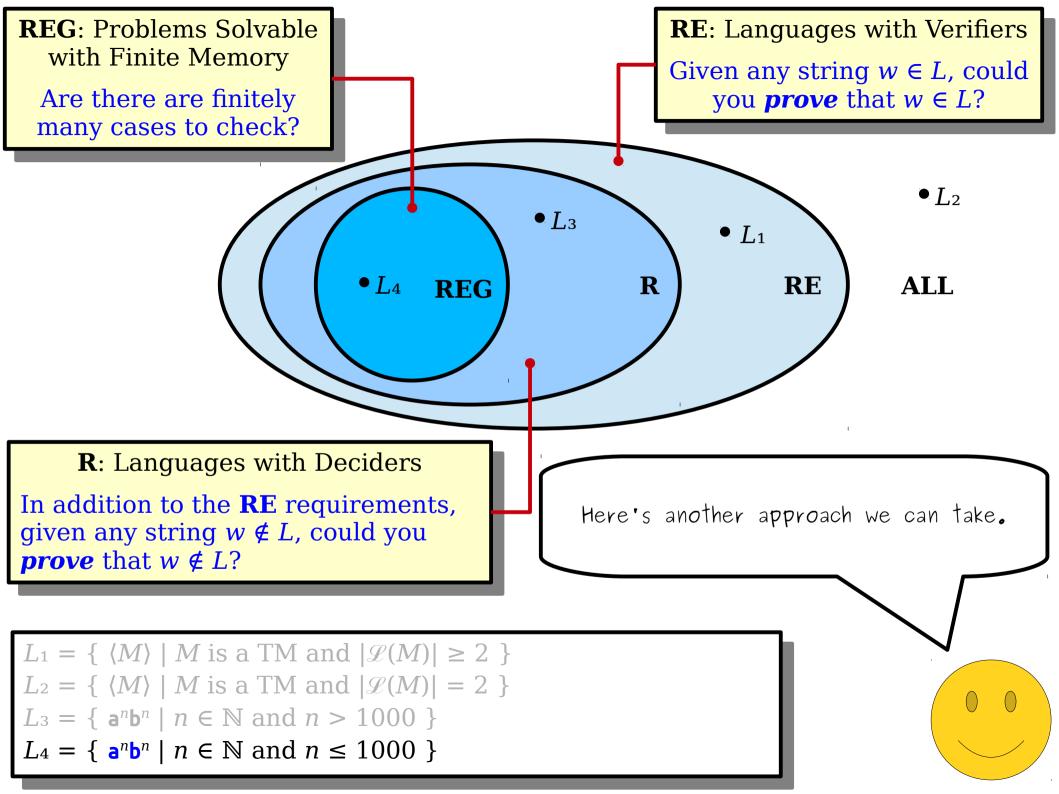


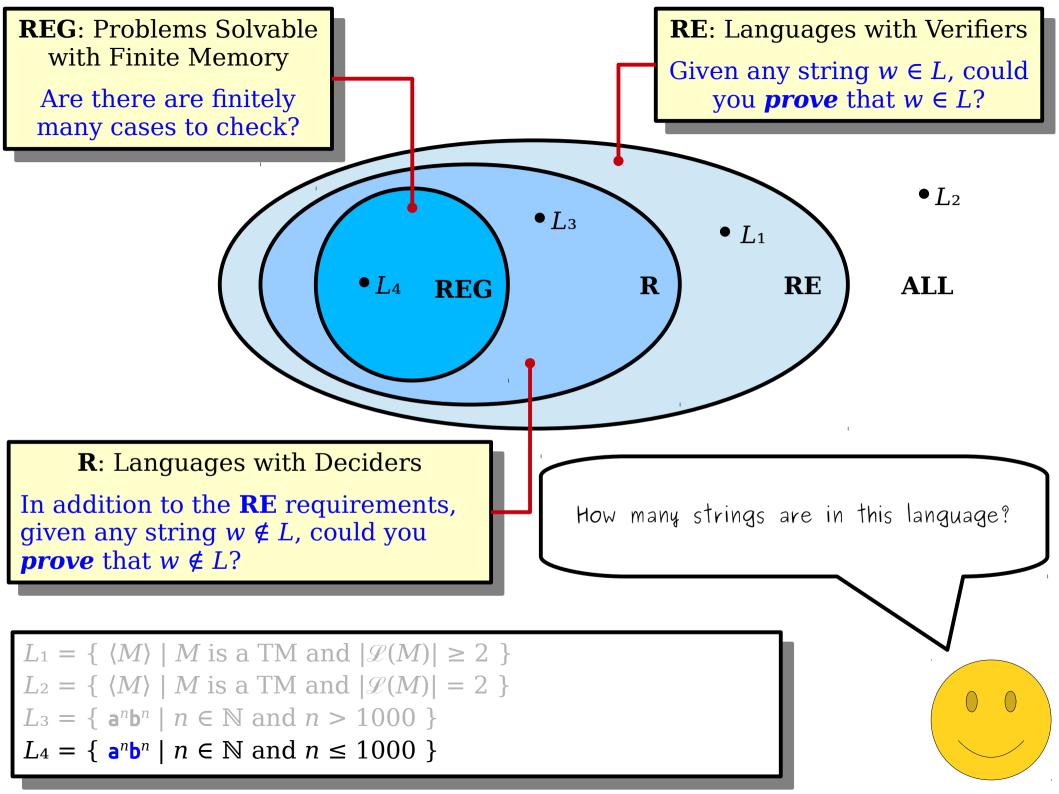


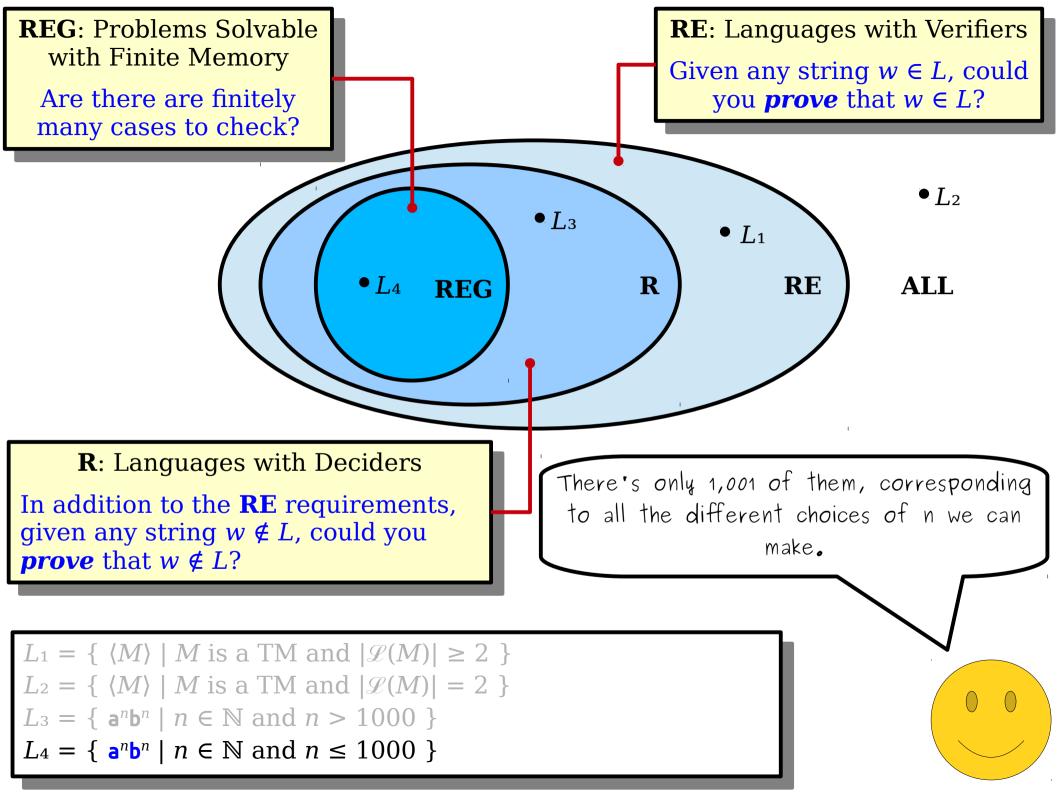


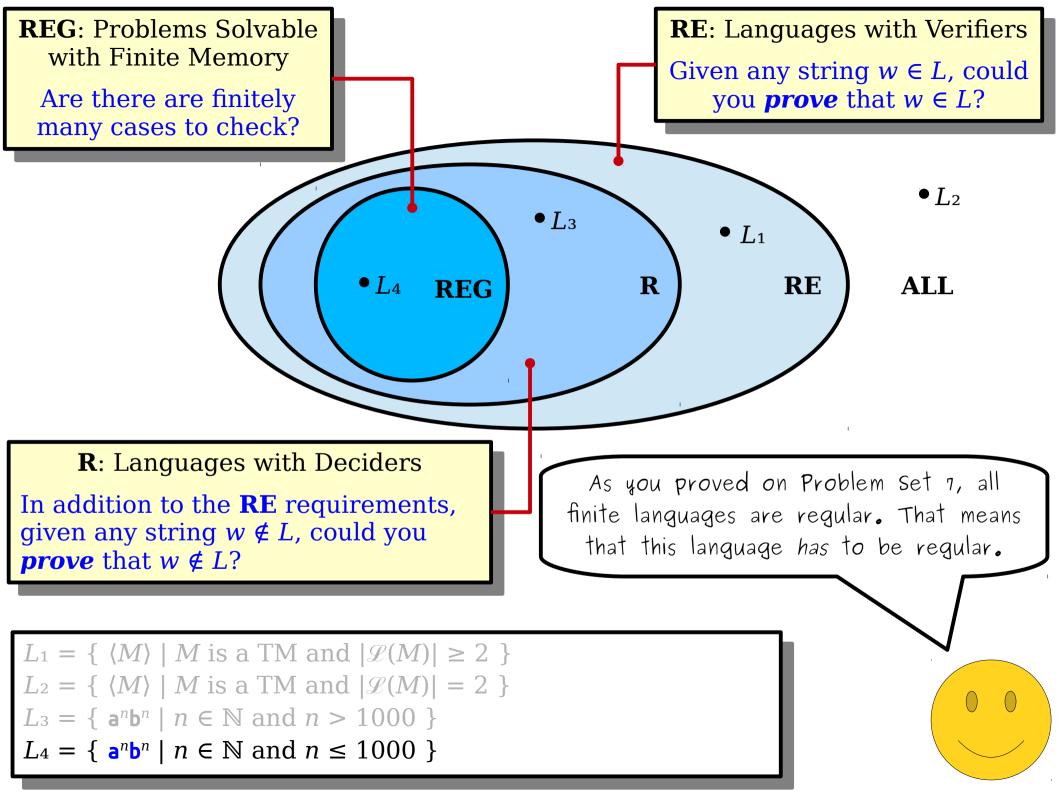


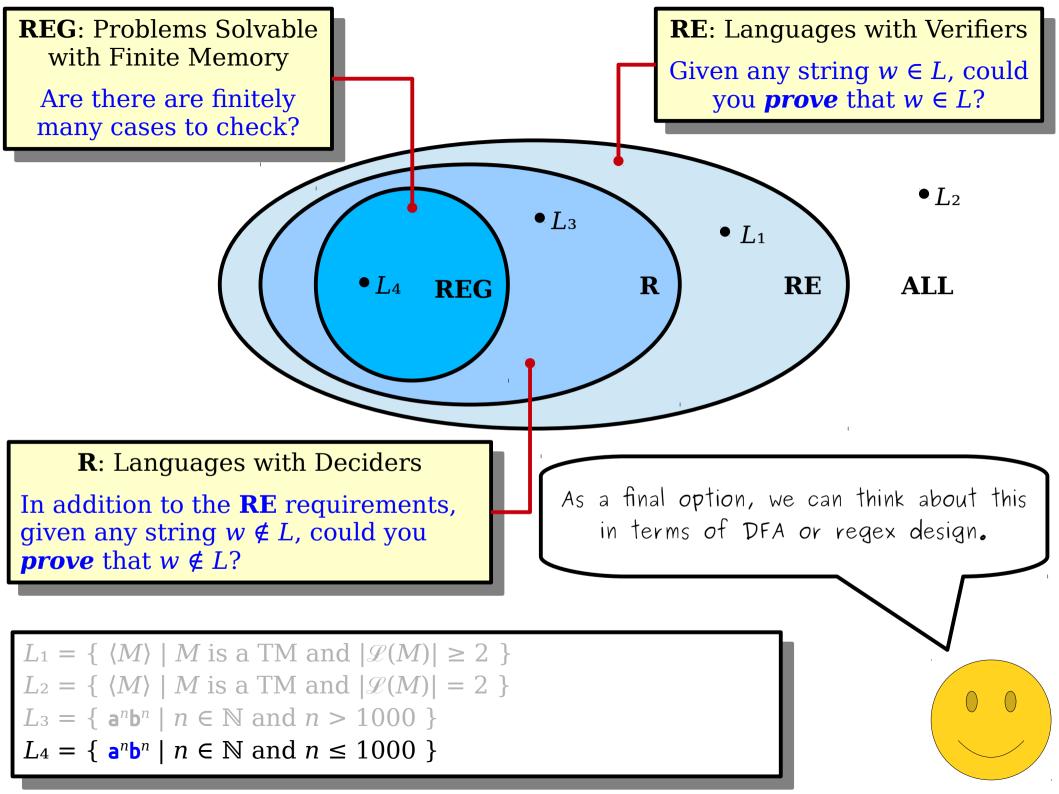


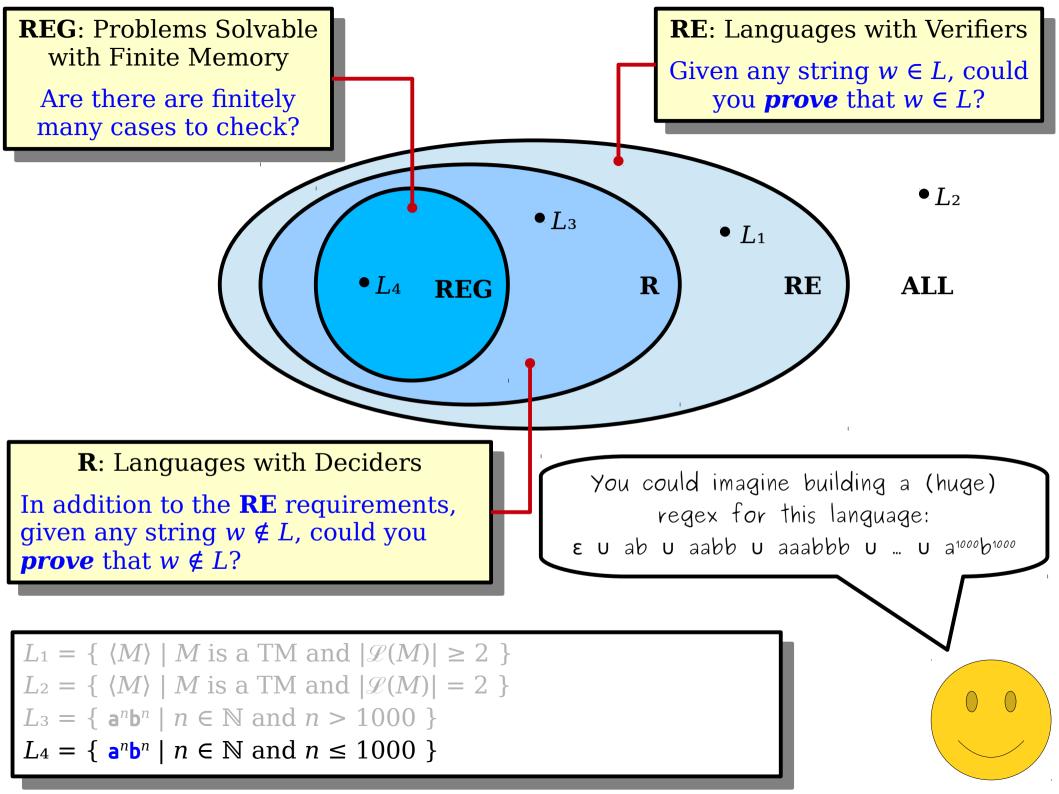


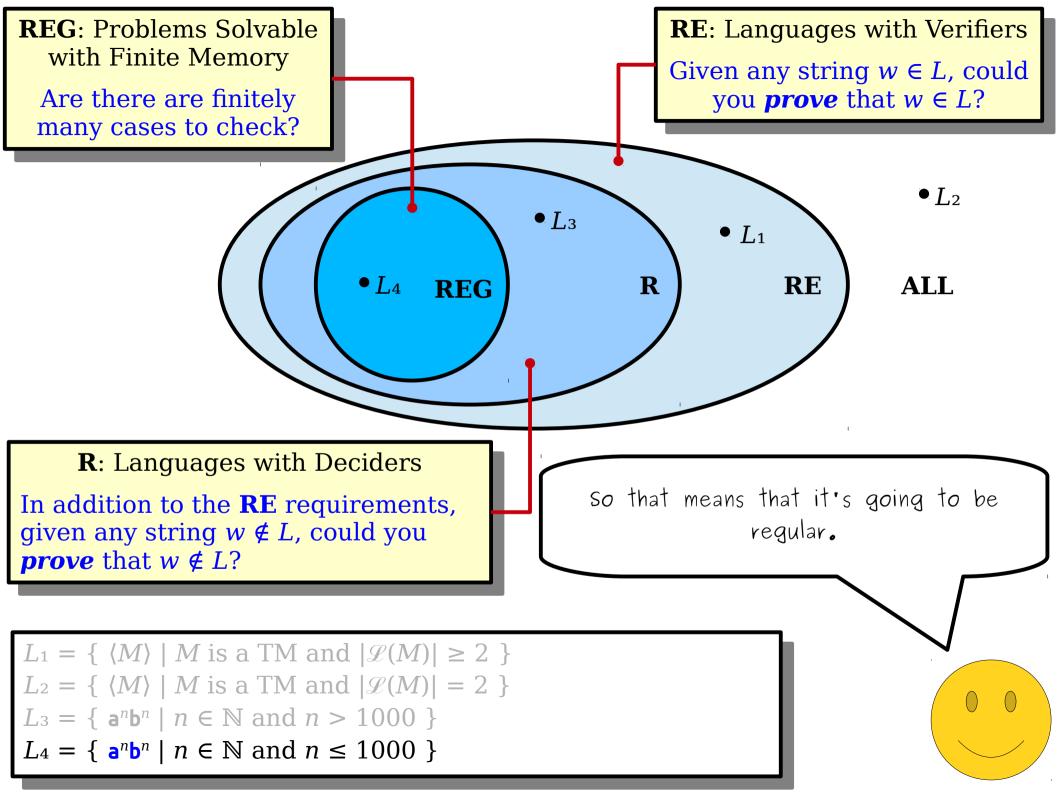


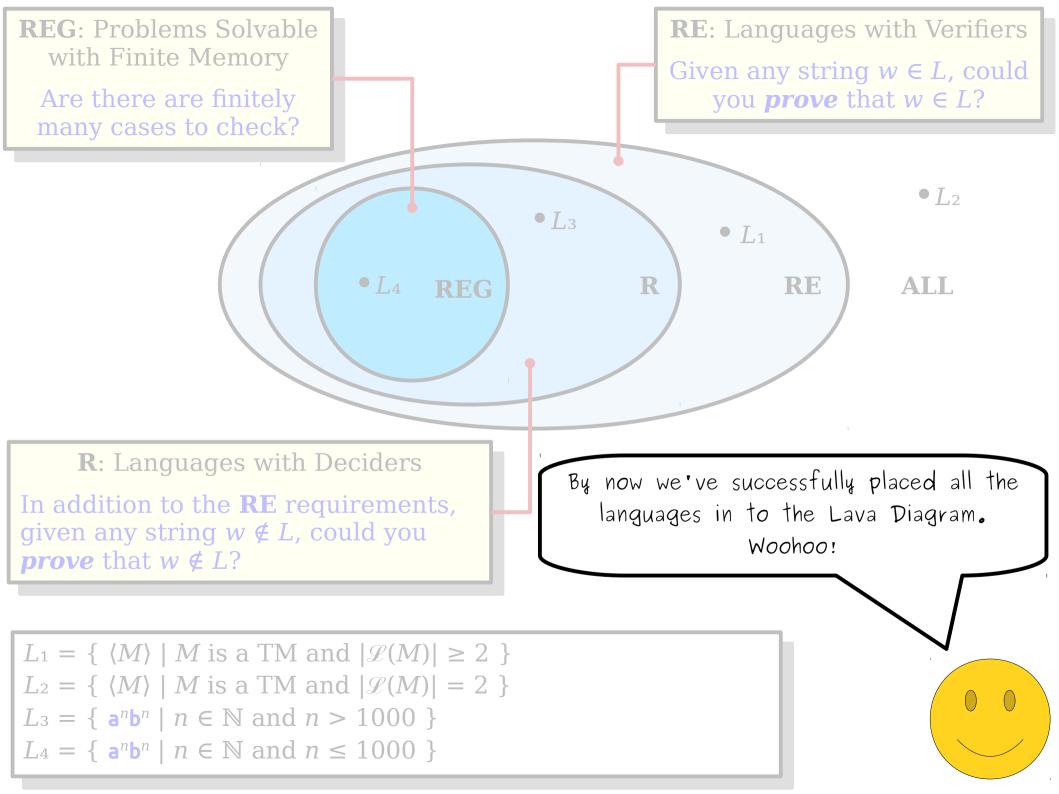


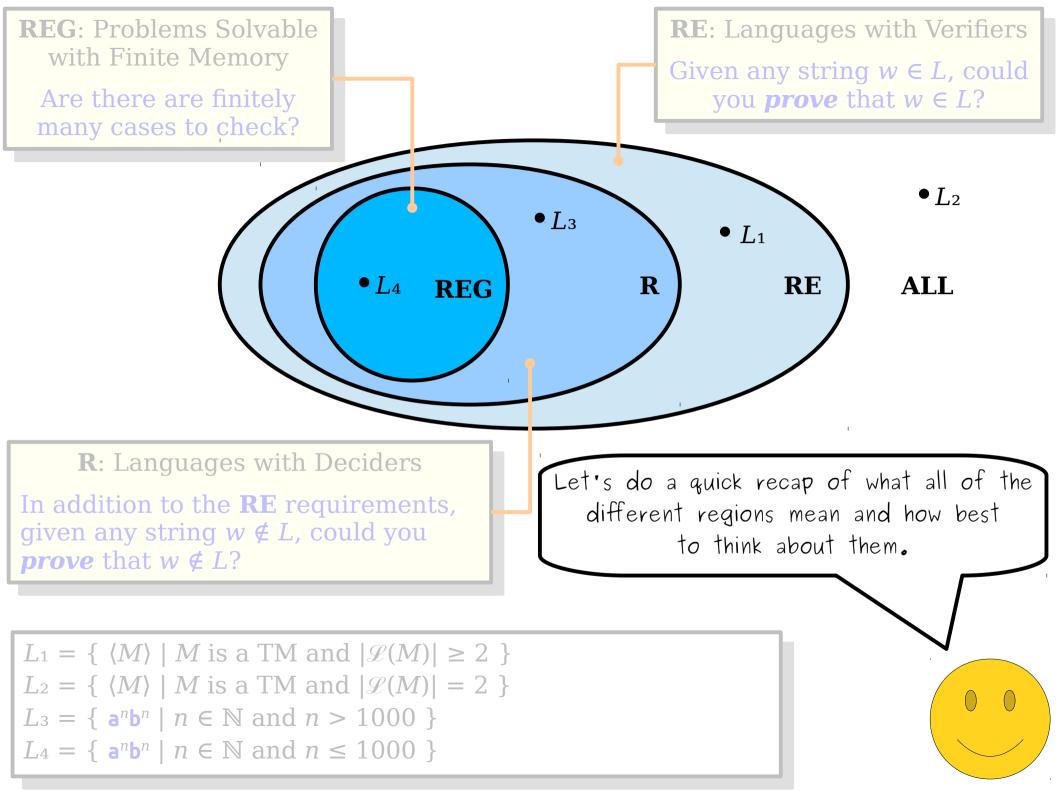


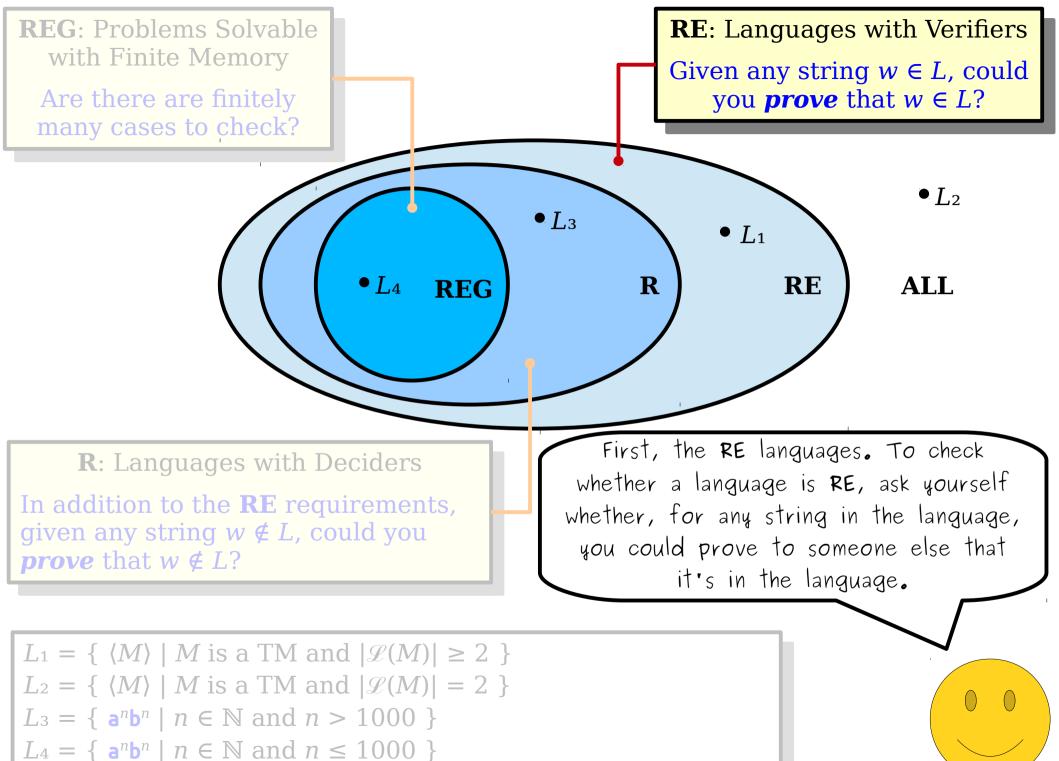


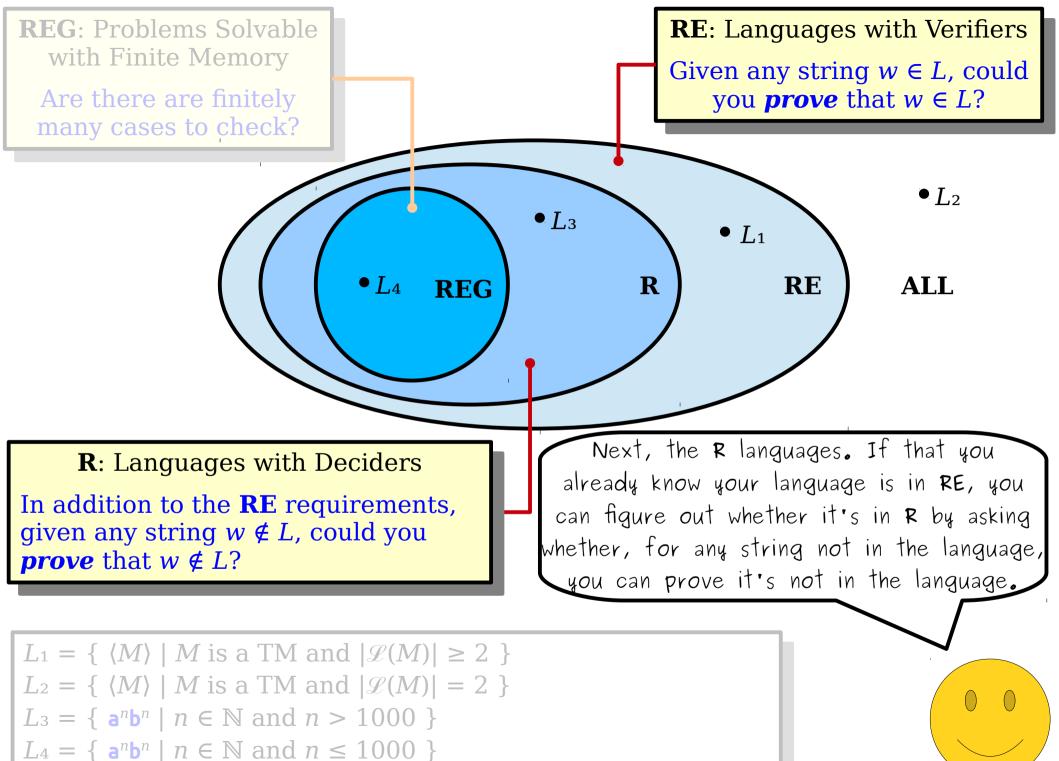


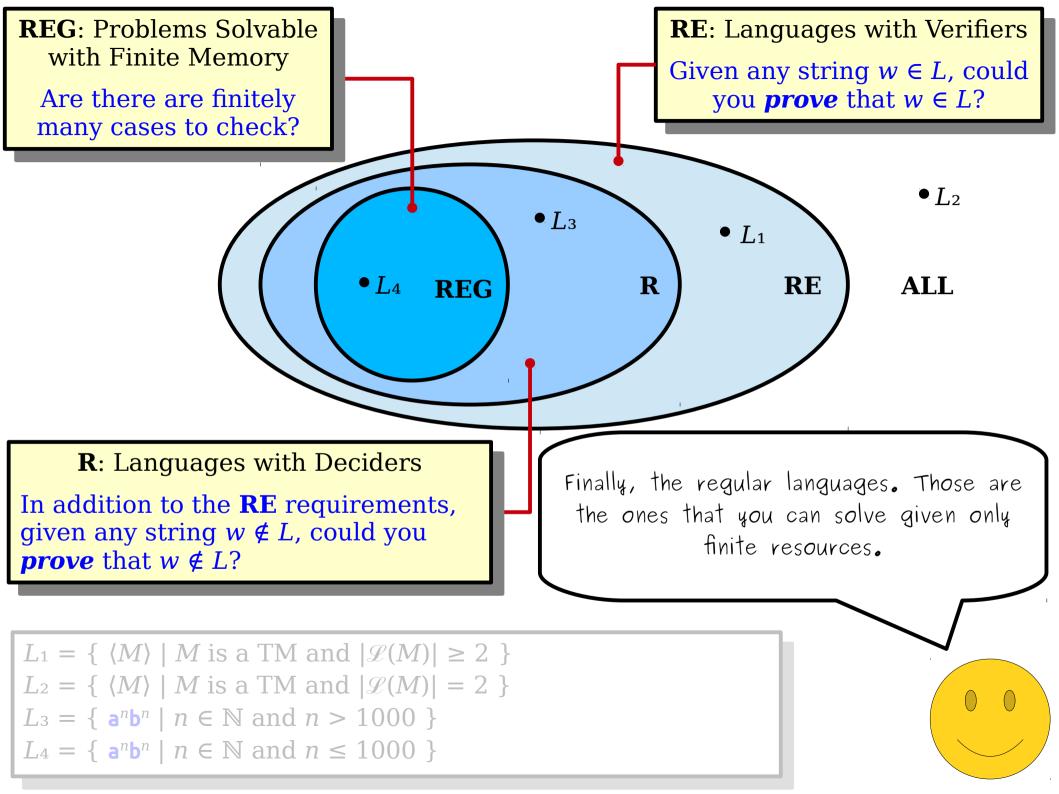


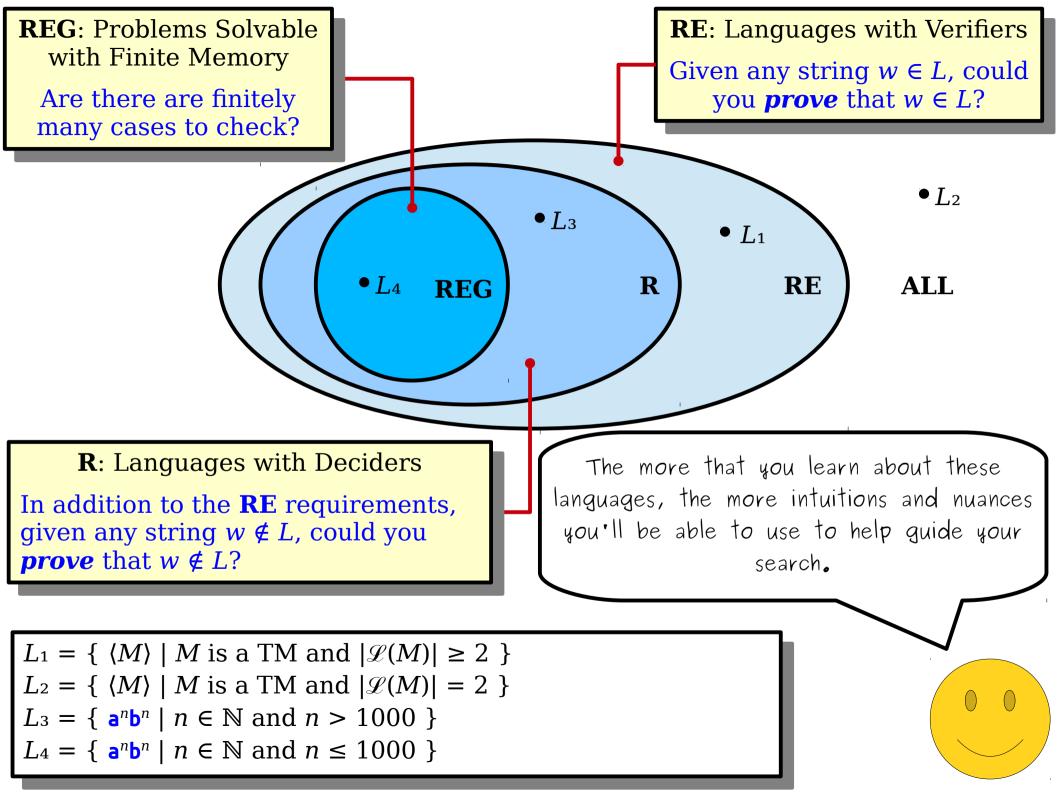


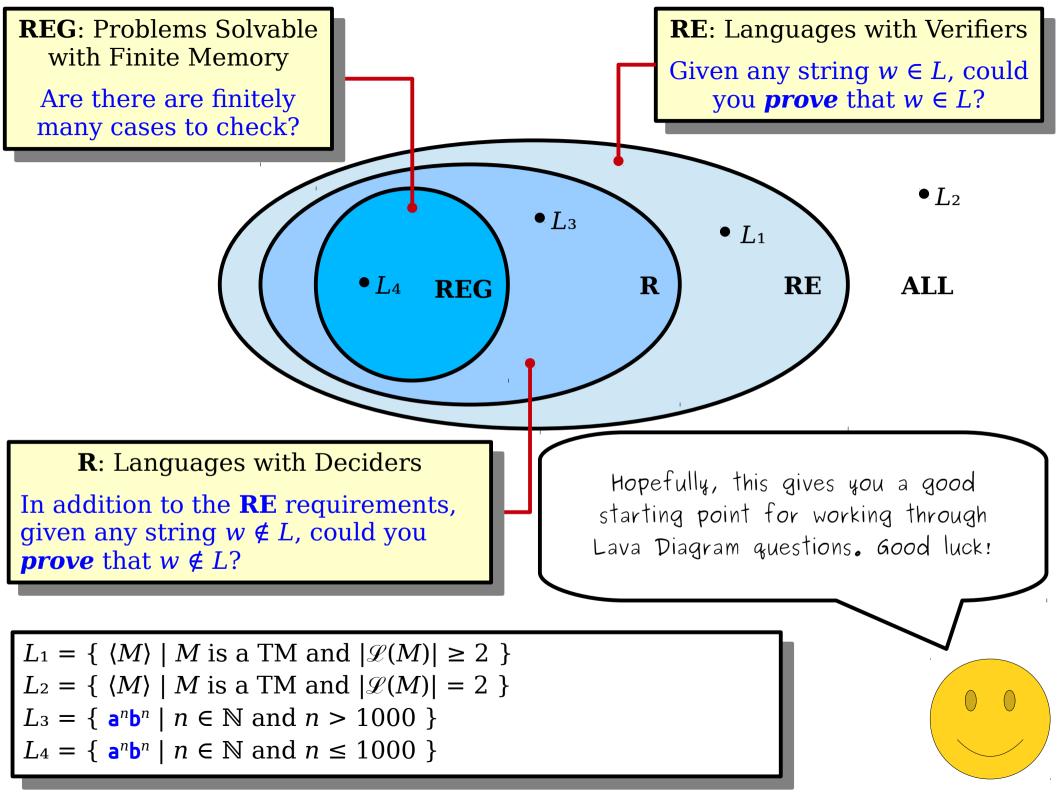












Hope this helps!

Please feel free to ask questions if you have them.

Did you find this useful? If so, let us know! We can go and make more guides like these.