

← [course home \(/table-of-contents\)](#)

## I'm making a search engine called MillionGazillion™.

I wrote a crawler that visits web pages, stores a few keywords in a database, and follows links to other web pages. I noticed that my crawler was wasting a lot of time visiting the same pages over and over, so I made a set, `visited`, where I'm storing URLs I've already visited. Now the crawler only visits a URL if it hasn't already been visited.

Thing is, the crawler is running on my old desktop computer in my parents' basement (where I totally don't live anymore), and it keeps running out of memory because `visited` is getting so huge.

How can I trim down the amount of space taken up by `visited`?

---

Want more coding interview help?

Check out **[interviewcake.com](https://www.interviewcake.com)** for more advice, guides, and practice questions.