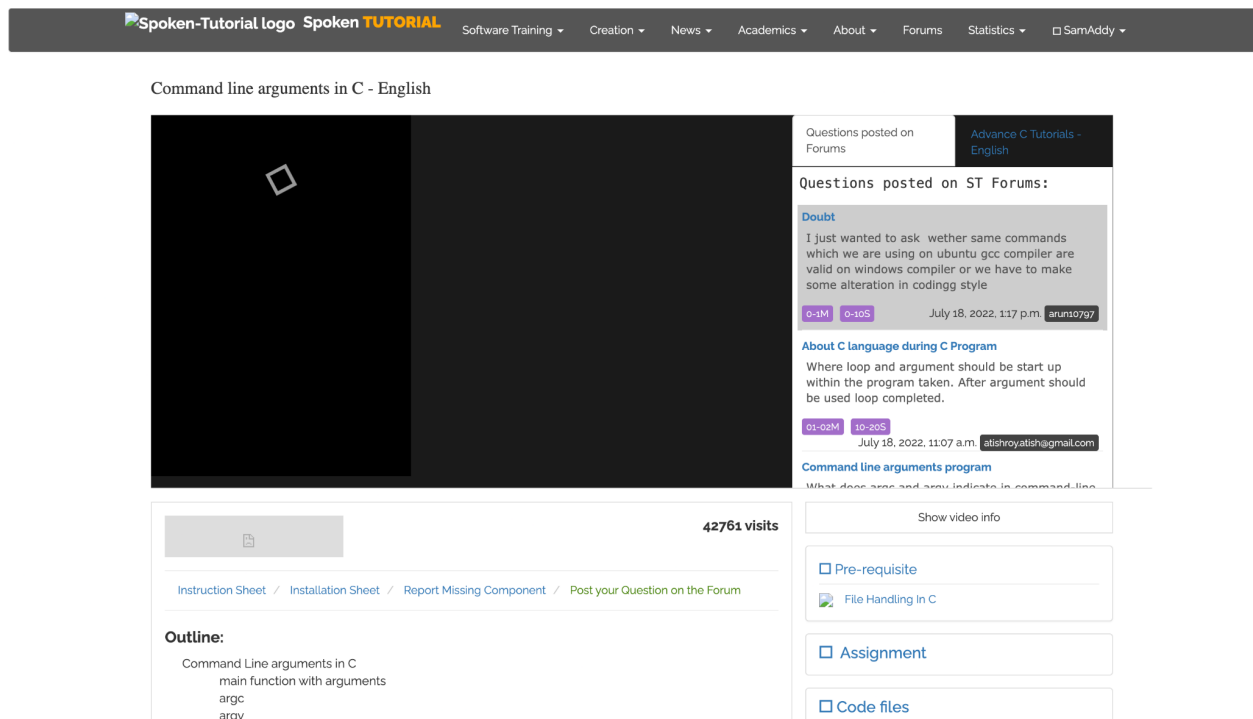


# Setup and updation guide for the mind map generator in a spoken tutorial

## Setting up :

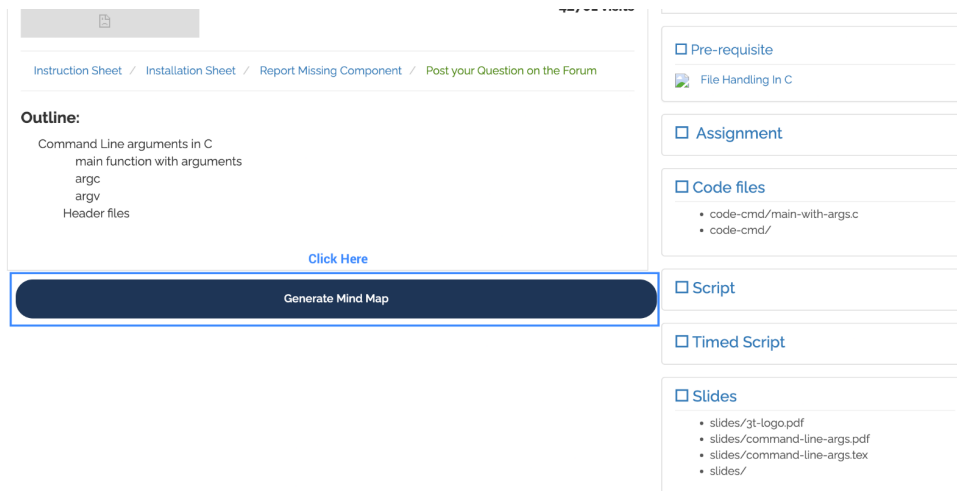
Please follow the [installation guide](#) to set up the project, start the Django server and open up the home page in the browser. It would look similar to the following,



The screenshot displays the Spoken Tutorial website interface. At the top, there is a navigation bar with the Spoken Tutorial logo and menu items: Software Training, Creation, News, Academics, About, Forums, Statistics, and SamAddy. The main content area is titled "Command line arguments in C - English". It features a video player on the left, which is currently black with a white play button icon. To the right of the video player, there is a forum section titled "Questions posted on ST Forums:". The forum contains two posts. The first post is titled "Doubt" and discusses the compatibility of C commands between Ubuntu GCC and Windows compilers. The second post is titled "About C language during C Program" and discusses the placement of loops and arguments within a program. Below the forum posts, there is a section for "Command line arguments program" with a video player. The video player has a "42761 visits" counter and a "Show video info" button. Below the video player, there is a navigation menu with links: "Instruction Sheet", "Installation Sheet", "Report Missing Component", and "Post your Question on the Forum". The "Outline:" section lists the topics covered in the video: "Command Line arguments in C", "main function with arguments", "argc", and "argv". On the right side of the page, there is a sidebar with a "Pre-requisite" section containing a link to "File Handling in C", an "Assignment" section, and a "Code files" section.

## Generate a mind map for the given spoken tutorial :

The code scraps the contents from the page linked to the Timed Script section to generate the mind map. Scroll down to the bottom of the page to see a button labelled "Generate Mind Map" and click it.



The screenshot shows a web page with a navigation bar at the top containing links: [Instruction Sheet](#), [Installation Sheet](#), [Report Missing Component](#), and [Post your Question on the Forum](#). Below the navigation bar is an **Outline:** section with the following text:

- Command Line arguments in C
  - main function with arguments
    - argc
    - argv
  - Header files

Below the outline is a [Click Here](#) link. At the bottom of the page is a large dark blue button labeled **Generate Mind Map**. On the right side of the page, there is a sidebar with several sections:

- Pre-requisite**: [File Handling In C](#)
- Assignment**
- Code files**:
  - [code-cmd/main-with-args.c](#)
  - [code-cmd/](#)
- Script**
- Timed Script**
- Slides**:
  - [slides/3t-logo.pdf](#)
  - [slides/command-line-args.pdf](#)
  - [slides/command-line-args.tex](#)
  - [slides/](#)

FOSSEE Project  
Scilab  
eSim (Oscad)  
Python  
OpenFoam (CFD)

Aakash Labs  
Co-learn  
Education Mission

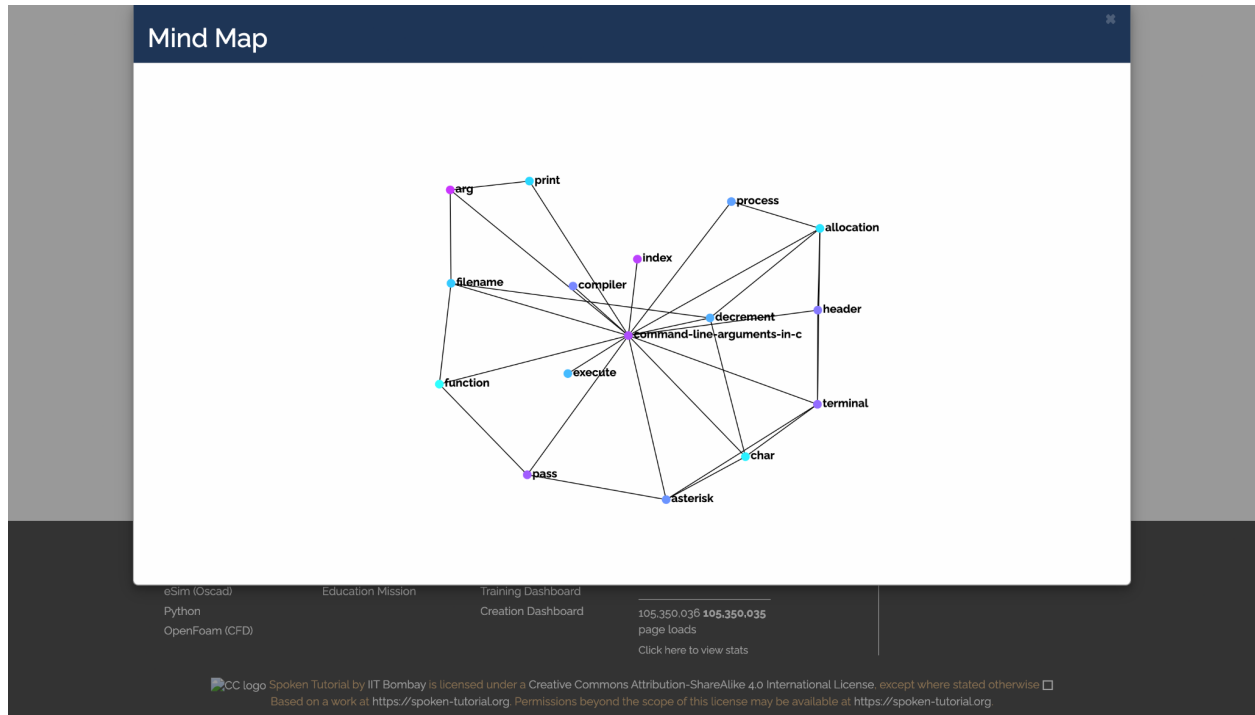
Software Training  
Contacts for Training  
Training Dashboard  
Creation Dashboard

About Us  
News  
105,350.036 **105,350,035**  
page loads  
[Click here to view stats](#)

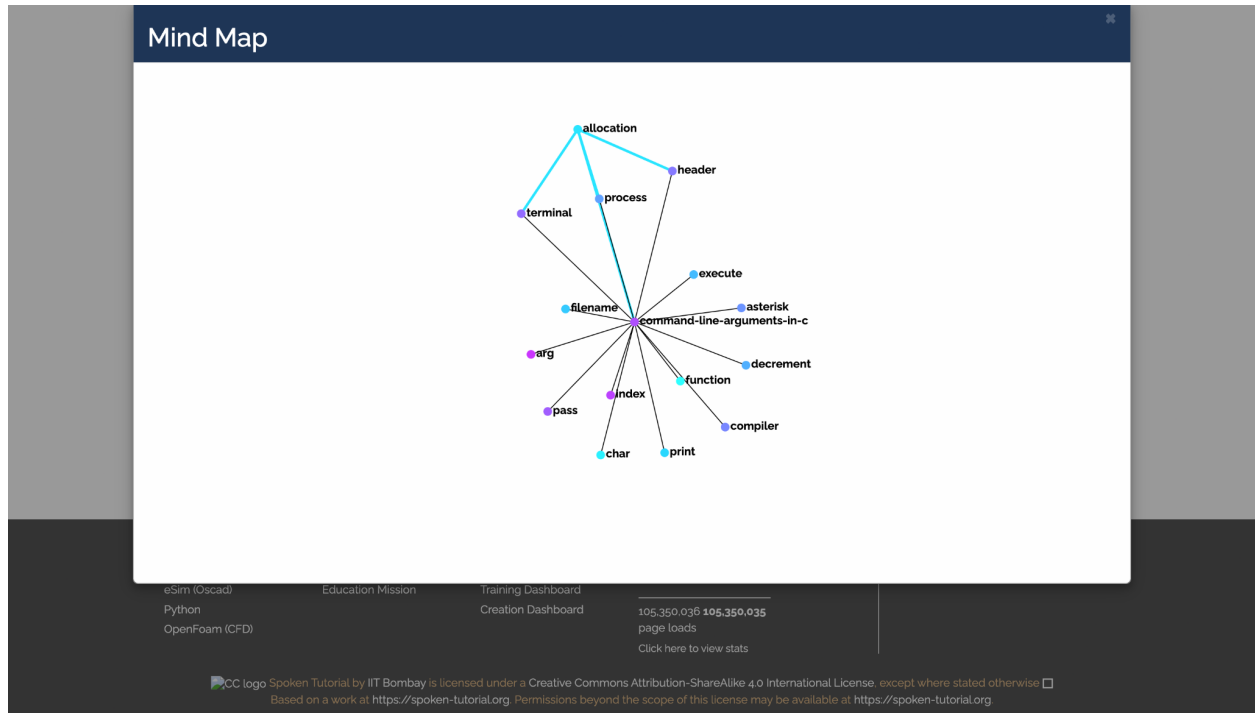
 ITB Logo  
Developed at IIT Bombay

## Redraw the mind map by emphasizing a particular keyword :

Clicking on the “Generate Mind Map” button will create a pop-up window, which would then display the mind map generated for that particular script in a few seconds. It would look similar to the following,



Further, based on the user's choice, they can select different keywords in the mind map by clicking to emphasize their relationships with other keywords, clicking a keyword would then redraw the mind map by highlighting those relationships as follows,



Users can close the pop-up using the **x** button at the top right.

## Updation

1. To avoid a particular word appearing in the mind map, it could be added to the [words\\_to\\_ignore.txt](#). This file contains all the highly frequent words that were collected from all timed scripts on the website which are supposedly less informative.
2. All contributor names are stored in [contributors.txt](#) to avoid names appearing in the mind map, any new names can be added there.
3. Three different [lemmatizers](#) could be used (NLTK, SpaCy and StanfordCoreNLP), and new ones can be added to the same file.
4. Two different [keyword extractors](#) could be used (yake, BERT), and new ones can be added to the same file.
5. [Word2Vec](#) is used for generating word embeddings, other models such as BERT, and GPT could be used instead as well. The process would just involve training those models and using the embeddings in the same way.
6. Changes to the threshold could be updated to tune the density of the graph in the same [word2vec.py](#) file.
7. The D3.js force simulation configuration could be updated through the [app.js](#) file.