



HAL
open science

Generative AI for Generative Programming: Automating Code Variants and Exploring the Boundaries of LLMs

Mathieu Acher

► **To cite this version:**

Mathieu Acher. Generative AI for Generative Programming: Automating Code Variants and Exploring the Boundaries of LLMs. 2024. hal-04809253

HAL Id: hal-04809253

<https://hal.science/hal-04809253v1>

Submitted on 28 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Generative AI for Generative Programming:

Automating Code Variants and Exploring the Boundaries of LLMs

Mathieu Acher



28 november 2024 @ Caen, invited talk at Normastic

<https://www.normastic.fr/event/journee-de-laxe-donnees-apprentissage-connaissances-large-language-models-llms/>



Large language models (LLMs) can be used to automate software engineering tasks, with the hope of boosting developers' productivity without sacrificing reliability.

In this talk, I will briefly present two initiatives: the défi Inria LLM4Code and the CodeCommons project around SoftwareHeritage, exemplifying ongoing research efforts.

Then, I will show how LLMs can automatically generate software variants across different technological spaces (Python, Rust, JavaScript, etc.) and implement new features from simple prompts.

I will also highlight how LLMs can assist in modernizing legacy applications, such as those written in COBOL or in deprecated technologies.

I will further discuss the concept of programming without directly manipulating programs, relying almost entirely on AI to manage technical details and customize code.

For all their potential, I will also demonstrate that LLMs can dramatically fail at times, making mistakes that human developers need to carefully verify and correct.

I will conclude with an illustration of LLM prompt sensitivity in chess, showcasing how LLMs can be used both to generate variants (e.g., prompts, hypotheses, ideas) and to accelerate scientific discovery through software-based exploration and analysis.

Plan

- Overview: LLM and software engineering tasks, LLM4Code, CodeCommons
- Generating software variants across different technological spaces (Python, Rust, JavaScript, etc.) and implement new features from simple prompts.
- Programming without directly manipulating programs, relying almost entirely on AI to manage technical details and create variants of code (customization).
- How LLMs can be used both to generate variants (e.g., prompts, hypotheses, ideas) and to accelerate scientific discovery through software-based exploration and analysis.
- Conclusion

LLM

PhD at University of Nice Sophia Antipolis (MESR, 2008-2011)

Postdoc at University of Namur, Belgium (Marie Curie fellowship, 2011-2012)

Associate Professor at University of Rennes 1/ISTIC (2012-2022)

Junior research fellow at IUF "mediation" (september 2021, 5 years)

Professor at INSA Rennes since september 2022

Love chess (FIDE master, 2341 Elo)

Love football (coach of Vétérans in Saint Aubin du Cormier)

Love software

Love science

can you depict a map of France with all locations?

```
python Copy code

import folium

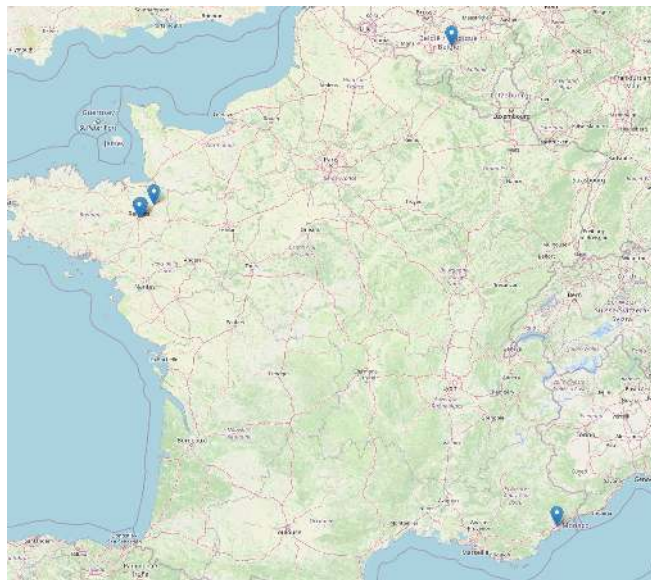
# Create a map centered around France
m = folium.Map(location=[46.603354, 1.888334], zoom_start=6)

# Add markers for each location
locations = {
    'University of Nice Sophia Antipolis': [43.6155, 7.0719],
    'University of Namur, Belgium': [50.4652, 4.8676],
    'University of Rennes 1/ISTIC': [48.1172, -1.6778],
    'INSA Rennes': [48.1215, -1.6366],
    'Saint Aubin du Cormier': [48.2972, -1.3746]
}

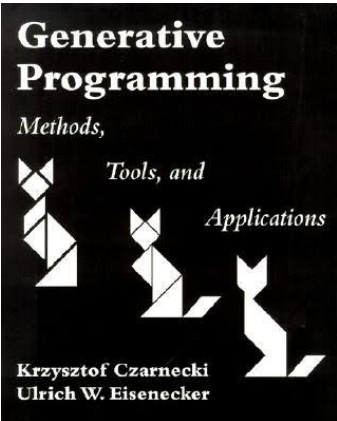
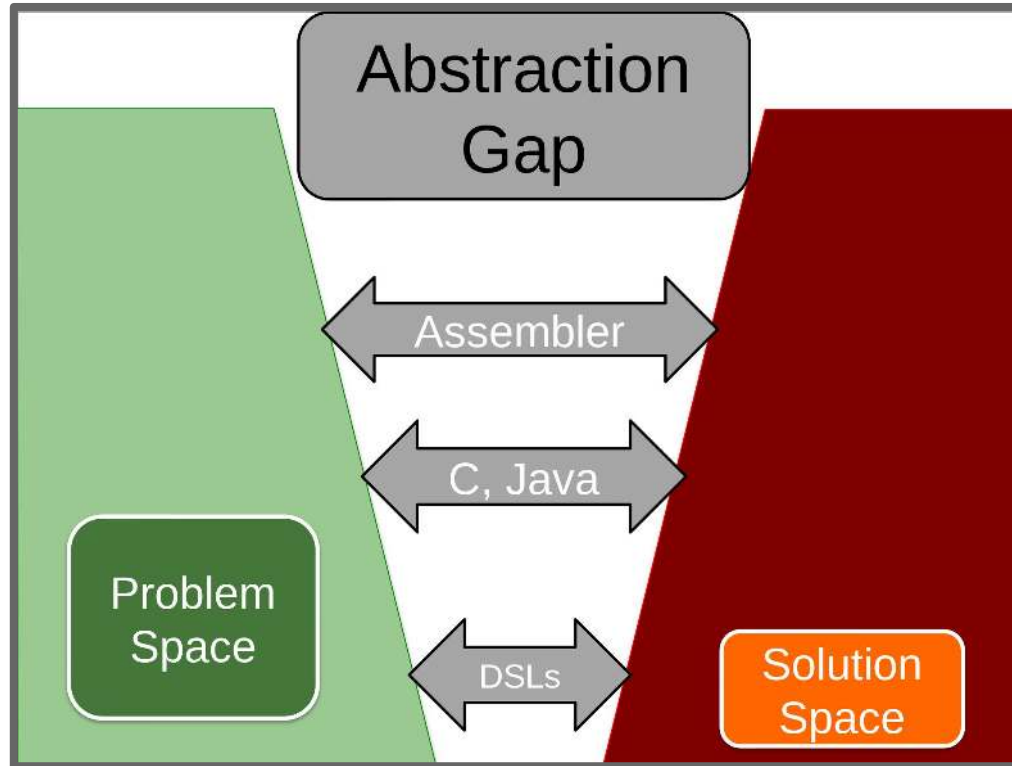
for place, coord in locations.items():
    folium.Marker(location=coord, popup=place).add_to(m)

# Save the map to an HTML file
m.save('map.html')
```

Software



Generative programming (Czarnecki and Eisenecker 2000), Software Product Line engineering: automatically generate software variants from a specification written in one or more textual or graphical domain-specific languages (DSLs)



Generative AI for Generative programming: automatically generate software variants from a specification written in one or more textual or graphical domain-specific languages (DSLs)



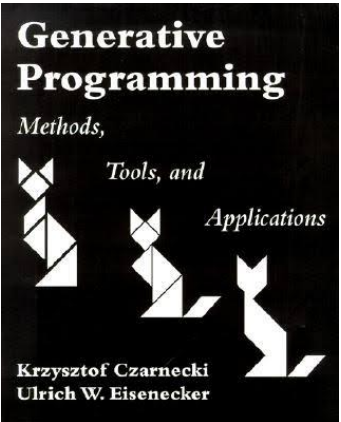
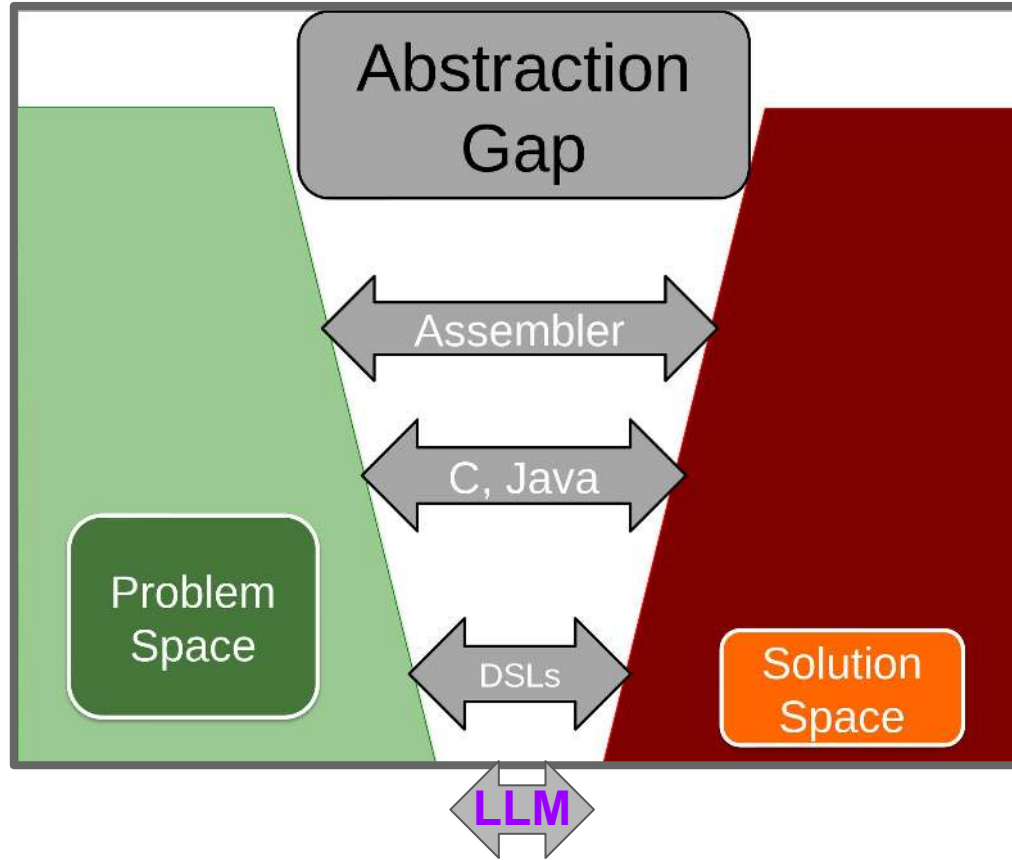
starcoder2 Public



AlphaCode

DeepSynth

GitHub Copilot



Reliable and productive code assistants based on LLMs

Main objectives:

- Generate better code: quality, diversification
- Validate the results: test, formal proof, review, integration
- Assist tedious tasks: boilerplate code, software migration

Organization: 3 cross-domain work packages

- **WP1:** Self-improving code generation
- **WP2:** Migration, maintenance and evolution of existing software
- **WP3:** Build interactive software development tools with AI-in-the-loop

LLM4Code: Reliable and Productive Code Assistants
Based on Large Language Models

Project Leads

DIVERSE	Rennes	Mathieu Acher	mathieu.acher@inria.fr
PICUBE	Paris	Guillaume Baudart	guillaume.baudart@inria.fr

Internal Partners (EP/CRI)

ARGO	Paris	Marc LeLARGE	marc.lelarge@inria.fr
DELI-CREAM	Socha	Anthony Bordg	anthony.bordg@inria.fr
FVREF	Lille	Anne-Filien	anne.etien@univ-lille.fr
FLOWERS	Bordeaux	Pierre-Yves Oudreyer	pierre-yves.oudreyer@inria.fr
MYTMO5VNE	Bordeaux	Xavier Hinaut	xavier.hinaut@inria.fr
SPIRALS	Lille	Clément Quinton	clement.quinton@inria.fr
STAMP	Sophia	Cyril Cohen	cyril.cohen@inria.fr
SED	Nancy	Laurent Pierrou	laurent.pierrou@inria.fr

External Partners

SOFTWARE HERITAGE	Roberto di Cosmo	roberto@dicosmo.org
LABRI	Nathanaël Fijalkow	nathanael.fijalkow@labri.fr
	Romain Robbes	romain.robbes@cnrs.fr
SOPRA STERIA	Yann Glorieu	yann.glorieu@sopraasteria.com

 You

DIVERSE	Rennes
PICUBE	Paris

Internal Partners (EP/C

ARGO	Paris
DEDUCTEAM	Saclay
EVREF	Lille
FLOWERS	Bordeaux
MNEMOSYNE	Bordeaux
SPIRALS	Lille
STAMP	Sophia
SED	Nancy

Please give me a downloadable map of France that depicts locations of these cities and teams

 ChatGPT

 Analyzing ▾

From picture to code SE task = code synthesis!

MA You

DIVERSE	Rennes
PICUBE	Paris

Internal Partners (EP/C

ARGO	Paris
DEDUCTEAM	Saclay
EVREF	Lille
FLOWERS	Bordeaux
MNEMOSYNE	Bordeaux
SPIRALS	Lille
STAMP	Sophia
SED	Nancy

Please give me a downloadable map of France that depicts locations of these cities and teams

ChatGPT

Analyzing ▾

python  Copy code

```
import folium

# Given coordinates for the cities (these are example coord
# Normally we would obtain these through a geocoding serv
coordinates = {
    "Rennes": [48.1173, -1.6778],
    "Paris": [48.8566, 2.3522],
    "Saclay": [48.7290, 2.1912],
    "Lille": [50.6292, 3.0573],
    "Bordeaux": [44.8378, -0.5792],
    "Sophia": [43.6222, 7.0500], # This is an approximat
    "Nancy": [48.6921, 6.1844]
}

# Initialize a map centered around France
map_france = folium.Map(location=[46.603354, 1.888334], zoom=10)

# Function to add a marker for a team
def add_team_marker(map_obj, team_name, city):
    folium.Marker(
        location=coordinates[city],
        popup=f"{team_name}, {city}",
        icon=folium.Icon(color='red', icon='info-sign')
    ).add_to(map_obj)

# Add markers for each team
add_team_marker(map_france, "DIVERSE", "Rennes")
add_team_marker(map_france, "PICUBE", "Paris")
add_team_marker(map_france, "ARGO", "Paris")
add_team_marker(map_france, "DEDUCTEAM", "Saclay")
add_team_marker(map_france, "EVEREF", "Lille")
add_team_marker(map_france, "FLOWERS", "Bordeaux")
add_team_marker(map_france, "MNEMOSYNE", "Bordeaux")
add_team_marker(map_france, "SPIRALS", "Lille")
add_team_marker(map_france, "STAMP", "Sophia")
add_team_marker(map_france, "SED", "Nancy")
```

From picture to app... SE task = code synthesis!

MA You

DIVERSE Rennes
PICUBE Paris

Internal Partners (EP/C

ARGO Paris
DEDUCTEAM Saclay
EVREF Lille
FLOWERS Bordeaux
MNEMOSYNE Bordeaux
SPIRALS Lille
STAMP Sophia
SED Nancy

Please give me a downloadable map and teams



ChatGPT

Analyzing ▾

python

Copy code

```
import folium

# Given coordinates for the cities (these are example coordinates)
# Normally we would obtain these through a geocoding service
coordinates = {
    "Rennes": [48.1173, -1.6778],
    "Paris": [48.8566, 2.3522],
    "Saclay": [48.7290, 2.1912],
    "Lille": [50.6292, 3.0573],
    "Bordeaux": [44.8378, -0.5792],
    "Sophia": [43.6222, 7.0500], # This is an approximate location
    "Nancy": [48.6921, 6.1844]
}

# Initialize a map centered around France
map_france = folium.Map(location=[46.603354, 1.888334], zoom_start=6)

# Function to add a marker for a team
def add_team_marker(map_obj, team_name, city):
    folium.Marker(
        location=coordinates[city],
        popup=f"{team_name}, {city}",
        icon=folium.Icon(color='red', icon='info-sign')
    ).add_to(map_obj)

# Add markers for each team
add_team_marker(map_france, "DIVERSE", "Rennes")
add_team_marker(map_france, "PICUBE", "Paris")
add_team_marker(map_france, "ARGO", "Paris")
add_team_marker(map_france, "DEDUCTEAM", "Saclay")
add_team_marker(map_france, "EVREF", "Lille")
add_team_marker(map_france, "FLOWERS", "Bordeaux")
add_team_marker(map_france, "MNEMOSYNE", "Bordeaux")
add_team_marker(map_france, "SPIRALS", "Lille")
add_team_marker(map_france, "STAMP", "Sophia")
add_team_marker(map_france, "SED", "Nancy")
```

Code synthesis: polyglot?

MA

You

I have the following Python code:

```
import folium
```

```
# Given coordinates for the cities (these are example coordinates and may not be accurate)
```

```
# Normally we would obtain these through a geocoding service
```

```
coordinates = {
```

```
    "Rennes": [48.1173, -1.6778],
```

```
    "Paris": [48.8566, 2.3522],
```

```
    "Saclay": [48.7290, 2.1912],
```

```
    "Lille": [50.6292, 3.0573],
```

```
    "Bordeaux": [44.8378, -0.5792],
```

```
    "Sophia": [43.6222, 7.0500], # This is an approximation for Sophia Antipolis
```

```
    "Nancy": [48.6921, 6.1844]
```

```
}
```

```
# Initialize a map centered around France
```

```
map_france = folium.Map(location=[46.603354, 1.888334], zoom_start=6)
```

```
# Add markers for each team
```

```
add_team_marker(map_france, "DIVERSE", "Rennes")
```

```
add_team_marker(map_france, "PICUBE", "Paris")
```

```
add_team_marker(map_france, "ARGO", "Paris")
```

```
add_team_marker(map_france, "DEDUCTEAM", "Saclay")
```

```
add_team_marker(map_france, "EVEREF", "Lille")
```

```
add_team_marker(map_france, "FLOWERS", "Bordeaux")
```

```
add_team_marker(map_france, "MNEMOSYNE", "Bordeaux")
```

```
add_team_marker(map_france, "SPIRALS", "Lille")
```

```
add_team_marker(map_france, "STAMP", "Sophia")
```

```
add_team_marker(map_france, "SED", "Nancy")
```

```
# Since we cannot display the folium map directly in this environment,
```

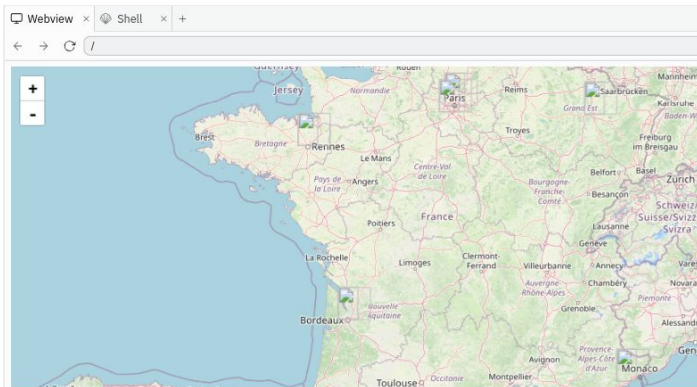
```
# we'll save it to an HTML file which you can then open in a web browser.
```

```
map_file_path = '/mnt/data/france_teams_map.html'
```

```
map_france.save(map_file_path)
```

```
map_file_path
```

Please write a JavaScript program, using Leaflet.js



```
script.js > f addCityMarker > X popupText
1 var map = L.map('map').setView([46.603354, 1.888334], 6);
2
3 // Add OpenStreetMap tiles:
4 L.tileLayer('https://{s}.tile.openstreetmap.org/{z}/{x}/{y}.png', {
5   maxZoom: 19,
6   attribution: '© OpenStreetMap contributors'
7 }).addTo(map);
8
9 // Teams and their respective cities
10 var teams = [
11   { city: "Rennes", team: "DIVERSE" },
12   { city: "Paris", team: "PICUBE" },
13   { city: "Paris", team: "ARGO" },
14   { city: "Saclay", team: "DEDUCTEAM" },
15   { city: "Lille", team: "EVEREF" },
16   { city: "Bordeaux", team: "FLOWERS" },
17   { city: "Bordeaux", team: "MNEMOSYNE" },
18   { city: "Lille", team: "SPIRALS" },
19   { city: "Sophia", team: "STAMP" },
20   { city: "Nancy", team: "SED" }
21 ];
22
23 // Helper function to get coordinates of a city
24 function getCoordinates(city) {
25   var coordinates = {
26     "Rennes": [48.1173, -1.6778],
27     "Paris": [48.8566, 2.3522],
28     "Saclay": [48.7290, 2.1912],
29     "Lille": [50.6292, 3.0573],
30     "Bordeaux": [44.8378, -0.5792],
31     "Sophia": [43.6222, 7.0500],
32     "Nancy": [48.6921, 6.1844]
33   };
34   return coordinates[city];
35 }
36
37 // Add markers for each city
38 object.keys(teamsByCity).forEach(function(city) {
39   addCityMarker(city, teamsByCity[city]);
40 });
41
42 }
```

ChatGPT

Creating a JavaScript program using Leaflet.js to achieve a similar functionality as your Python code with Folium is a good way to visualize geographic data on web platforms. Leaflet is a leading open-source JavaScript library for mobile-friendly interactive maps.

Here's a basic example of how you could write such a program:

SE task: modernization/migration



You

I have the following Python code:

```
import folium
```

```
# Given coordinates for the cities (these are example coordinates and may not be accurate)
```

```
# Normally we would obtain these through a geocoding service
```

```
coordinates = {
```

```
    "Rennes": [48.1173, -1.6778],
```

```
    "Paris": [48.8566, 2.3522],
```

```
    "Saclay": [48.7290, 2.1912],
```

```
    "Lille": [50.6292, 3.0573],
```

```
    "Bordeaux": [44.8378, -0.5792],
```

```
    "Sophia": [43.6222, 7.0500], # This is an approximation for Sophia Antipolis
```

```
    "Nancy": [48.6921, 6.1844]
```

```
}
```

```
# Initialize a map centered around France
```

```
map_france = folium.Map(location=[46.603354, 1.888334], zoom_start=6)
```

```
# Add markers for each team
```

```
add_team_marker(map_france, "DIVERSE", "Rennes")
```

```
add_team_marker(map_france, "PICUBE", "Paris")
```

```
add_team_marker(map_france, "ARGO", "Paris")
```

```
add_team_marker(map_france, "DEDUCTEAM", "Saclay")
```

```
add_team_marker(map_france, "EVEREF", "Lille")
```

```
add_team_marker(map_france, "FLOWERS", "Bordeaux")
```

```
add_team_marker(map_france, "MNEMOSYNE", "Bordeaux")
```

```
add_team_marker(map_france, "SPIRALS", "Lille")
```

```
add_team_marker(map_france, "STAMP", "Sophia")
```

```
add_team_marker(map_france, "SED", "Nancy")
```

```
# Since we cannot display the folium map directly in this environment,
```

```
# we'll save it to an HTML file which you can then open in a web browser.
```

```
map_file_path = '/mnt/data/france_teams_map.html'
```

```
map_france.save(map_file_path)
```

```
map_file_path
```

Please write a JavaScript program, using Leaflet.js

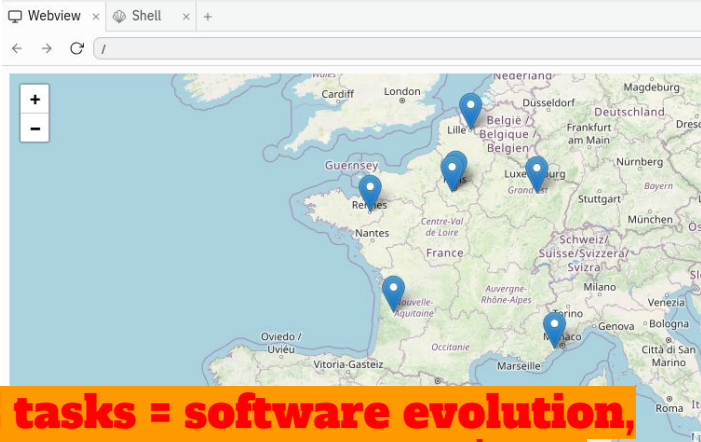


ChatGPT

Creating a JavaScript program using Leaflet.js to achieve a similar functionality as your Python code with Folium is a good way to visualize geographic data on web platforms. Leaflet is a leading open-source JavaScript library for mobile-friendly interactive maps.

Here's a basic example of how you could write such a program:

Software evolves...



SE tasks = software evolution, maintenance, upgrade | bug fixing | program understanding

```

index.html x  script.js x  +
script.js > f addCityMarker > X popupText
1  var map = L.map('map').setView([46.603354, 1.888334], 6);
2
index.html x  script.js x  +
index.html
1  <!DOCTYPE html>
2  <html>
3
4  <head>
5  <title>Team Locations in France</title>
6  <link rel="stylesheet" href="https://unpkg.com/leaflet@1.1.0/dist/leaflet.css" />
7  <script src="https://unpkg.com/leaflet@1.1.0/dist/leaflet.js"></script>
8  </head>
9  <style>
10 #map {
11   height: 400px;
12 }
13 </style>
14
15 <body>
16 <div id="map"></div>
17 <script src="script.js"></script>
18 </body>
19
20
21
22
23
24
25
26
27
28
29 10   height: 400px;
30 11   }
31 12 </style>
32 13 </head>
33 14
34 15 <body>
35 16 <div id="map"></div>
36 17 <script src="script.js"></script>
37 18 </body>
38 19
39 20 </html>
40
41 // Helper function to get coordinates of a city
42 function getCoordinates(city) {
43   var coordinates = {
44     "Rennes": [48.1173, -1.6778],
45     "Paris": [48.8566, 2.3522],
46     "Saclay": [48.7290, 2.1912],
47     "Lille": [50.6292, 3.0573],
48     "Bordeaux": [44.8378, -0.5792],
49     "Sophia": [43.6222, 7.0500],
50     "Nancy": [48.6921, 6.1844]
51   };
52   return coordinates[city];
53 }
54
55 // Add markers for each city
56 Object.keys(teamsByCity).forEach(function(city) {
57   addCityMarker(city, teamsByCity[city]);
58 });

```


<https://arxiv.org/abs/2210.14699>

Piloting Copilot and Codex: Hot Temperature, Cold Prompts, or Black Magic?

```
def choose_num(x, y):  
    """  
    This function takes two positive  
    numbers x and y and returns the  
    biggest even integer number that  
    is in the range [x, y] inclusive.  
    If there's no such number, then  
    the function should return -1.  
  
    For example:  
    choose_num(12, 15) = 14  
    choose_num(13, 12) = -1  
    """  
    if x % 2 != 0:  
        x += 1  
    if y % 2 != 0:  
        y -= 1  
    if x > y:  
        return -1  
    if x % 2 == 0:  
        x += 1  
    while x <= y:  
        if x % 2 == 0:  
            return x  
        x += 2  
    return -1
```

(a) Original prompt

```
def choose_num(x, y):  
    """  
    This function takes two positive  
    numbers x and y and returns the  
    biggest even integer number that  
    is in the range [x, y] inclusive.  
    If there's no such number, then  
    the function should return -1.  
  
    For example:  
    choose_num(12, 15) = 14  
    choose_num(13, 12) = -1  
  
    Write a quick algorithm to  
    solve this problem.  
    """  
    if x % 2 == 0:  
        x = x  
    else:  
        x = x + 1  
    if y % 2 == 0:  
        y = y  
    else:  
        y = y - 1  
    if x > y:  
        return -1  
    else:  
        return max(x, y)
```

(b) Modified prompt 1

```
def choose_num(x, y):  
    """  
    This function takes two positive  
    numbers x and y and returns the  
    biggest even integer number that  
    is in the range [x, y] inclusive.  
    If there's no such number, then  
    the function should return -1.  
    """  
    if x % 2 == 1:  
        x += 1  
    if y % 2 == 0:  
        y -= 1  
    for i in range(x, y + 1, 2):  
        if x <= i <= y:  
            return i  
    return -1
```

(c) Modified prompt 2

Jean-Baptiste Döderlein
ENS Rennes
Rennes, France
jean-baptiste.doderlein@ens-rennes.fr

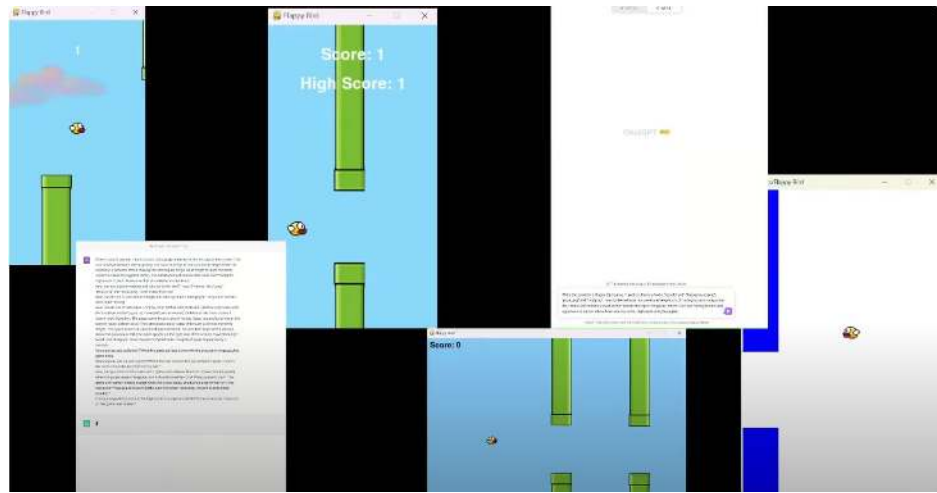
Djamel Eddine Khelladi
CNRS, Univ Rennes, IRISA, Inria
Rennes, France
djamel-eddine.khelladi@irisa.fr

Mathieu Acher
Univ Rennes, IUF, CNRS, Inria, IRISA
Rennes, France
mathieu.acher@irisa.fr

Benoit Combemale
Univ Rennes, CNRS, Inria, IRISA
Rennes, France
benoit.combemale@irisa.fr



Prompt 1: “Write a Flappy Bird game in python.”



Prompt 42: “Create a Flappy Bird game using Pygame. The game window should have a width of 288 and a height of 512, with the caption “Flappy Bird”. The game should include the following elements:

1. Load the game assets: a background image (“assets/background.png”), a bird image (“assets/bird.png”), a pipe image (“assets/pipe.png”), and a cloud image (“assets/cloud.png”).

....

10. Quit the game and close the Pygame window when the game loop is exited. Make sure to use the provided images and code structure to create a functional Flappy Bird game.”

https://www.youtube.com/watch?v=nmx1pg_DHzw

https://dev.to/diverse_research/end-user-programming-of-flappy-bird-with-chatgpt-a-reality-check-18lk

here: SE ~ end-user software engineering

Reliable and productive code assistants based on LLMs

Main objectives:

- Generate better code: quality, diversification
- Validate the results: test, formal proof, review, integration
- Assist tedious tasks: boilerplate code, software migration

Organization: 3 cross-domain work packages

- **WP1:** Self-improving code generation
- **WP2:** Migration, maintenance and evolution of existing software
- **WP3:** Build interactive software development tools with AI-in-the-loop

<https://project.inria.fr/llm4code/>

LLM4Code: Reliable and Productive Code Assistants
Based on Large Language Models

Project Leads

DIVERSE	Rennes	Mathieu Acher	mathieu.acher@inria.fr
PICUBE	Paris	Guillaume Baudart	guillaume.baudart@inria.fr

Internal Partners (EP/CRI)

ARGO	Paris	Marc LeLARGE	marc.lelarge@inria.fr
DELI-CREAM	Socha	Anthony Bordg	anthony.bordg@inria.fr
FUTUR	Lille	Anne-Françoise	anne.stien@univ-lille.fr
FLOWERS	Bordeaux	Pierre-Yves Oudeyer	pierre-yves.oudeyer@inria.fr
MYTOS/VNE	Bordeaux	Xavier Hinaut	xavier.hinaut@inria.fr
SPIRALS	Lille	Clément Quinton	clement.quinton@inria.fr
STAMP	Sophia	Cyril Cohen	cyril.cohen@inria.fr
SED	Nancy	Laurent Pierrou	laurent.pierrou@inria.fr

External Partners

SOFTWARE HERITAGE	Roberto di Cosmo	roberto@dicosmo.org
LABRI	Nathanaël Fijalkov	nathanael.fijalkov@labri.fr
	Romain Robbes	romain.robbes@cnrs.fr
SOPRA STERIA	Yann Glorieu	yann.glorieu@sopraasteria.com

Software/code: huge data....

Key to the “success” of code LLMs: training over such data (caveats: private/ad-hoc data; lack of traceability; many legal and ethical issues; quality matters)

Cultural Heritage Industry Research Public Administration



Source files

16,786,947,392



Commits

3,555,307,942



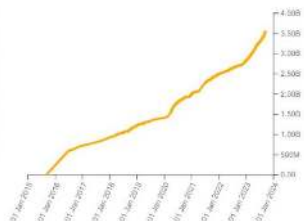
Projects

262,077,475



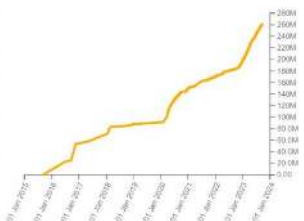
Directories

13,486,743,792



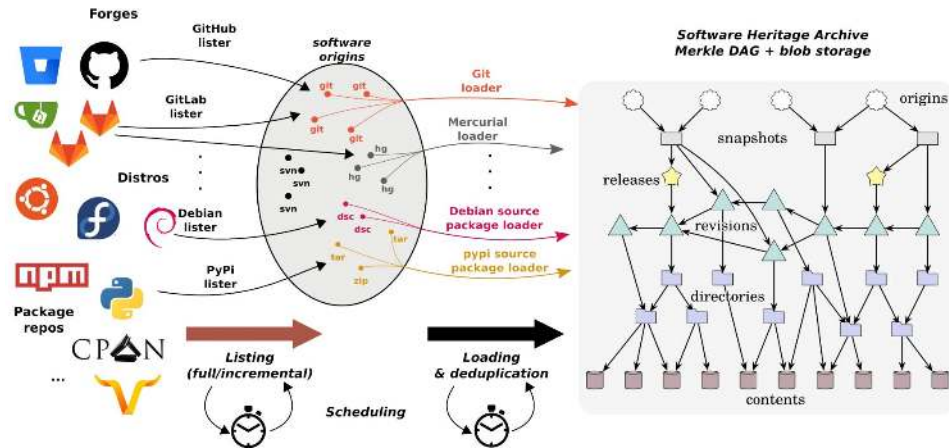
Authors

63,835,635



Releases

74,477,682

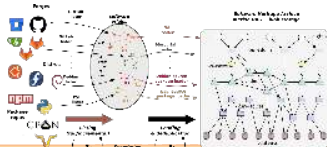




CODE COMMONS: “LE” BIEN COMMUN FOR SOURCE CODE IN FRANCE!



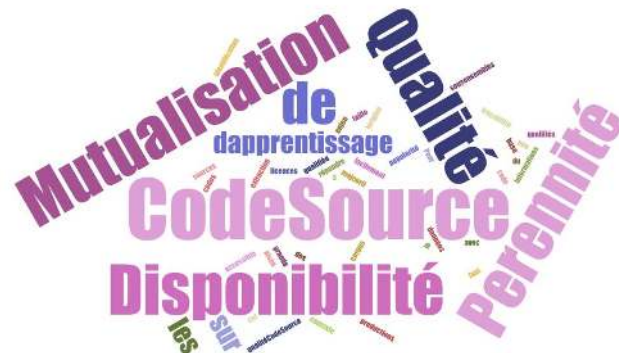
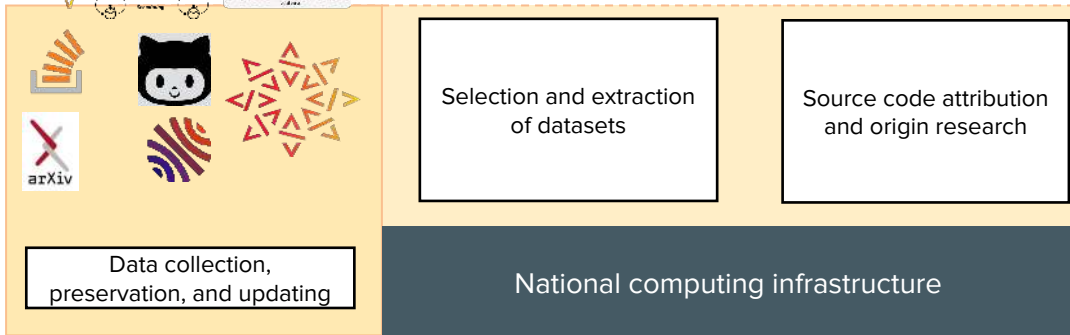
.....



Generative AI models



User group



For new generation AI!



starcoder2

Public

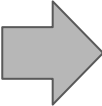
Generating software variants across different technological spaces (Python, Rust, JavaScript, etc.) and implement new features from simple prompts.



16 variants of Rust programs

Write a configurable **Rust** program that implements four features and their selections, namely, Hello, Beautiful, Wonderful, and World... If the feature is selected, then we print the name of the feature. For instance, we could well print "Hello Beautiful World", "Hello World", "Hello", "World" etc. The four features should be implemented as compile-time features.

(LLM) [...] you can use the 'cfg' attribute to enable or disable the features at compile-time [...] in your 'Cargo.toml' file [...].

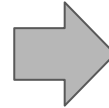


```
fn main() {  
    #[cfg(feature = "hello")]  
    print!("Hello ");  
    #[cfg(feature = "beautiful")]  
    print!("Beautiful ");  
    #[cfg(feature = "wonderful")]  
    print!("Wonderful ");  
    #[cfg(feature = "world")]  
    print!("World");  
    println!();  
}
```

16 variants of Java programs

Write a configurable Java program that implements [...] The four features should be implemented as compile-time features. As there is no preprocessor and built-in support for conditional compilation in Java, it is needed to leverage a templating engine for customization, acting as a preprocessor for Java. You can use Jinja (Python)

(LLM) To create a configurable Java program that implements the four features as compile-time features using Jinja (Python) templating engine, you need to follow these steps: ... Create a Jinja template (main_template.java.j2) for your Java program [...] ... Create a Python script (generate_java.py) that will generate the Java source file based on the provided features:...



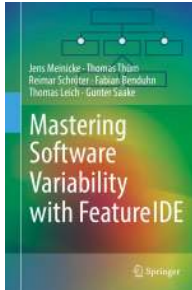
```
public class Main {
    public static void main(String[] args) {
{% if features.Hello %}
        System.out.print("Hello ");
{% endif %}
{% if features.Beautiful %}
        System.out.print("Beautiful ");
{% endif %}
{% if features.Wonderful %}
        System.out.print("Wonderful ");
{% endif %}
{% if features.World %}
        System.out.print("World");
{% endif %}
        System.out.println();
    }
}
```

```
100          P Conditional Compilation with FeatureIDE
1 public class Main {
2     public static void main(String[] args) {
3         System.out.print("Hello ");
4         System.out.print("Beautiful ");
5         System.out.print("Wonderful ");
6         System.out.print("World");
7     }
8 }
9
10 public class Main {
11     public static void main(String[] args) {
12         System.out.print("Hello ");
13         System.out.print("Beautiful ");
14         System.out.print("Wonderful ");
15         System.out.print("World");
16     }
17 }
18 }
```

Fig. 9.1 Java World example using the Jinja preprocessor. Generated source file on the left-hand side and generated file on the right

if [features.Hello], if [features.Beautiful], if [features.Wonderful], and if [features.World] as shown in the left-hand side of Fig. 9.1. These conditions establish a mapping between features and source code, specifying whether the lines between these directives will be included into the generated source file. For the Hello World example, the preprocessor can be executed as follows:

```
Usage: -DHello -DBeautiful -DWorld Main.java targetDir
```



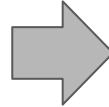
https://github.com/acherm/progvary-withgpt/blob/main/conditioncompilation-helloworld/ChatGPT-Java_program_with_Jinja.md

16 variants of Java programs and a generator

Write a configurable Java program that implements [...] The four features should be implemented as compile-time features. As there is no preprocessor and built-in support for conditional compilation in Java, it is needed to leverage a templating engine for customization, acting as a preprocessor for Java. You can use Jinja (Python)

(LLM) To create a configurable Java program that implements the four features as compile-time features using Jinja (Python) templating engine, you need to follow these steps: ... Create a Jinja template (main_template.java.j2) for your Java program [...] ... Create a Python script (generate_java.py) that will generate the Java source file based on the provided features:...

```
public class Main {
    public static void main(String[] args) {
{% if features.Hello %}
        System.out.print("Hello ");
{% endif %}
{% if features.Beautiful %}
        System.out.print("Beautiful ");
{% endif %}
{% if features.Wonderful %}
        System.out.print("Wonderful ");
{% endif %}
{% if features.World %}
        System.out.print("World");
{% endif %}
        System.out.println(
    }
}
```



```
import sys
from jinja2 import Template

def generate_java_code(features):
    with open('main_template.java.j2', 'r') as template_file:
        template = Template(template_file.read())
        java_code = template.render(features=features)
    return java_code

def write_java_code(java_code, output_file):
    with open(output_file, 'w') as java_file:
        java_file.write(java_code)

if __name__ == '__main__':
    features = {
        "Hello": True,
        "Beautiful": False,
        "Wonderful": True,
        "World": True
    }

    java_code = generate_java_code(features)
    output_file = 'Main.java'
    write_java_code(java_code, output_file)
    print(f"Java code generated in {output_file}")
```

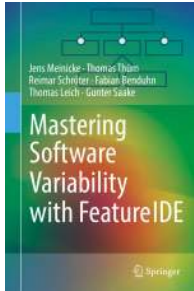
```
100 P Conditional Compilation with FeatureIDE
1 public class Main {
2     public static void main(String[]
3         args) {
4         System.out.print("Hello ");
5         // ...
6         System.out.print("Beautiful");
7         // ...
8         System.out.print("Wonderful");
9         // ...
10        System.out.print("World");
11        // ...
12        System.out.println("World");
13    }
14 }
```

Fig. 9.1 Java World example using the 'Simple preprocessor' annotation scheme like on the left-hand side and generated file on the right

if [features.Hello], if [features.Beautiful], and if [features.Wonderful] as shown in the left-hand side of Fig. 9.1. These conditions establish a mapping between features and source code, specifying whether the lines between these directives will be included into the generated source file. For the Hello World example, the preprocessor can be executed as follows:

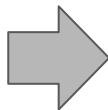
```
Usage: java -DHello -DBeautiful -DWorld Main.java targetDir
```

https://github.com/acherm/progvary-withgpt/blob/main/conditioncompilation-helloworld/ChatGPT-Java_program_with_Jinja.md



16 variants of JavaScript programs

Write a configurable JavaScript program that implements [...]
The four features should be implemented as feature toggles. I'd like to use Unleash, a feature toggle service, integrated with Gitlab



```
const { initialize } = require('unleash-client');

const unleash = initialize({
  appName: 'my-app-name',
  instanceId: 'my-instance-id',
  url: 'https://unleash.example.com/api/',
});

function printFeatures() {
  let output = '';

  if (unleash.isEnabled('hello')) {
    output += 'Hello';
  }

  if (unleash.isEnabled('beautiful')) {
    output += (output ? ' ' : '') + 'Beautiful';
  }

  if (unleash.isEnabled('wonderful')) {
    output += (output ? ' ' : '') + 'Wonderful';
  }

  if (unleash.isEnabled('world')) {
    output += (output ? ' ' : '') + 'World';
  }

  console.log(output);
}

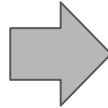
unleash.on('ready', printFeatures);
unleash.on('update', printFeatures);
```

https://github.com/acherm/progvary-withgpt/blob/main/featuretoggle-helloworld/ChatGPT-Unleash_JS_Feature_Toggles.md

16 variants of JavaScript programs at runtime

Write a configurable JavaScript program that implements [...]
The four features should be implemented as feature toggles. I'd like to use Unleash, a feature toggle service, integrated with Gitlab

“would it be possible to adapt the program as a Web application... and provide a server-side that returns the output while the client side prints it”



```
const express = require('express');
const { initialize } = require('unleash-client');
const cors = require('cors');

const app = express();
const unleash = initialize({
  appName: 'my-app-name',
  instanceId: 'my-instance-id',
  url: 'https://unleash.example.com/api/',
});

app.use(cors());

app.get('/features', (req, res) => {
  let output = '';

  if (unleash.isEnabled('hello')) {
    output += 'Hello';
  }

  if (unleash.isEnabled('beautiful')) {
    output += (output ? ' ' : '') + 'Beautiful';
  }

  if (unleash.isEnabled('wonderful')) {
    output += (output ? ' ' : '') + 'Wonderful';
  }

  if (unleash.isEnabled('world')) {
    output += (output ? ' ' : '') + 'World';
  }

  res.json({ output });
});

const PORT = process.env.
app.listen(PORT, () => {
  console.log('Server run
});

<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>Feature Toggles Web App</title>
</head>
<body>
  <h1>Enabled Features:</h1>
  <div id="output"></div>

  <script>
    async function fetchFeatures() {
      const response = await fetch('http://localhost:3000/features');
      const data = await response.json();
      document.getElementById('output').innerText = data.output;
    }

    fetchFeatures();
  </script>
</body>
</html>
```

https://github.com/acherm/progvary-withgpt/blob/main/featuretoggle-helloworld/ChatGPT-Unleash_JS_Feature_Toggles.md

3 variants of C programs

I have the following program. I'd like to create variations over this program using conditional compilation. Specifically, I'd like to create a first variation that consists in replacing the type long double by double... and another alternative variation that consists in replacing long double by float.

```
#include <stdio.h>
int main() {
    // Declare and initialize variables
    long double x = 1.0;
    long double y = 0x1p-53 + 0x1p-64;
    long double z = x + y;

    // Print the value of z
    printf("z = %a\n", z);
    return 0;
}
```



```
#include <stdio.h>

// Define the type to use as an alias
#ifdef USE_DOUBLE
    typedef double real;
#elif defined(USE_FLOAT)
    typedef float real;
#else
    typedef long double real;
#endif

int main() {
    // Declare and initialize variables
    real x = 1.0;
    real y = 0x1p-53 + 0x1p-64;
    real z = x + y;

    // Print the value of z
#ifdef USE_DOUBLE
    printf("z (using double) = %a\n", z);
#elif defined(USE_FLOAT)
    printf("z (using float) = %a\n", z);
#else
    printf("z (using long double) = %La\n", z);
#endif

    return 0;
}
```

```
#include <stdio.h>

// Define the type to use as an alias
#ifdef USE_DOUBLE
    typedef double real;
#elif defined(USE_FLOAT)
    typedef float real;
#else
    typedef long double real;
#endif

int main() {
    // Declare and initialize variables
    real x = 1.8;
    real y = 0x1p-53 + 0x1p-64;
    real z = x * y;

    // Print the value of z
#ifdef USE_DOUBLE
    printf("z (using double) = %f\n", z);
#elif defined(USE_FLOAT)
    printf("z (using float) = %f\n", z);
#else
    printf("z (using long double) = %Lf\n", z);
#endif

    return 0;
}
```

Write a Python script that compiles and executes the three variants above with and without the `-mfpmath=387` of `gcc`... Compile also with `clang`. Synthesize all results in a CSV containing the configuration (flag, with or without `-mfpmath=387`, compiler) and the different outputs. Please use `pandas` for writing the CSV... also the fieldnames should be `compiler`, `flags`, `type`, and `result`. The C file to compile is `assoc.c`

We then refine and ask:

Can you think about `gcc` flags or `clang` flags that could be combined/used and have an influence on floating point operations? Please implement them in the Python code.



```
def compile_and_run(compiler, flags):
    commands = [
        f"{compiler} {flags} -o original approx.c",
        f"{compiler} {flags} -DUSE_DOUBLE -o double_variation approx.c",
        f"{compiler} {flags} -DUSE_FLOAT -o float_variation approx.c",
    ]

    results = []
    for cmd in commands:
        subprocess.run(cmd, shell=True, check=True)

        binary = cmd.split()[-2]
        output = subprocess.check_output(f'./{binary}', shell=True).decode().strip()
        results.append(output)

    return results

def main():
    configurations = [
        ('gcc', ''),
        ('gcc', '-mfpmath=387'),
        ('gcc', '-ffast-math'),
        ('gcc', '-fno-signed-zeros'),
        ('gcc', '-fno-trapping-math'),
        ('gcc', '-fno-rounding-math'),
        ('clang', ''),
        ('clang', '-mno-sse'),
        ('clang', '-ffast-math'),
        ('clang', '-fno-signed-zeros'),
        ('clang', '-fno-trapping-math'),
        ('clang', '-fno-rounding-math'),
    ]

    data = []

    for compiler, flags in configurations:
        results = compile_and_run(compiler, flags)
        data.extend([
            {'compiler': compiler, 'flags': flags, 'type': 'original', 'result': results[0]},
            {'compiler': compiler, 'flags': flags, 'type': 'double', 'result': results[1]},
            {'compiler': compiler, 'flags': flags, 'type': 'float', 'result': results[2]},
        ])

    return data
```

Lots of variants to build programs!

```

def compile_and_run(compiler, flags):
    commands = [
        f"{compiler} {flags} -o original approx.c",
        f"{compiler} {flags} -DUSE_DOUBLE -o double_variation approx.c",
        f"{compiler} {flags} -DUSE_FLOAT -o float_variation approx.c",
    ]

    results = []
    for cmd in commands:
        subprocess.run(cmd, shell=True, check=True)

        binary = cmd.split()[-2]
        output = subprocess.check_output(f'./{binary}', shell=True).decode().strip()
        results.append(output)

    return results

def main():
    configurations = [
        ('gcc', ''),
        ('gcc', '-mfpmath=387'),
        ('gcc', '-ffast-math'),
        ('gcc', '-fno-signed-zeros'),
        ('gcc', '-fno-trapping-math'),
        ('gcc', '-fno-rounding-math'),
        ('clang', ''),
        ('clang', '-mno-sse'),
        ('clang', '-ffast-math'),
        ('clang', '-fno-signed-zeros'),
        ('clang', '-fno-trapping-math'),
        ('clang', '-fno-rounding-math'),
    ]

    data = []

    for compiler, flags in configurations:
        results = compile_and_run(compiler, flags)
        data.extend([
            {'compiler': compiler, 'flags': flags, 'type': 'original', 'result': results[0]},
            {'compiler': compiler, 'flags': flags, 'type': 'double', 'result': results[1]},
            {'compiler': compiler, 'flags': flags, 'type': 'float', 'result': results[2]},
        ])

```

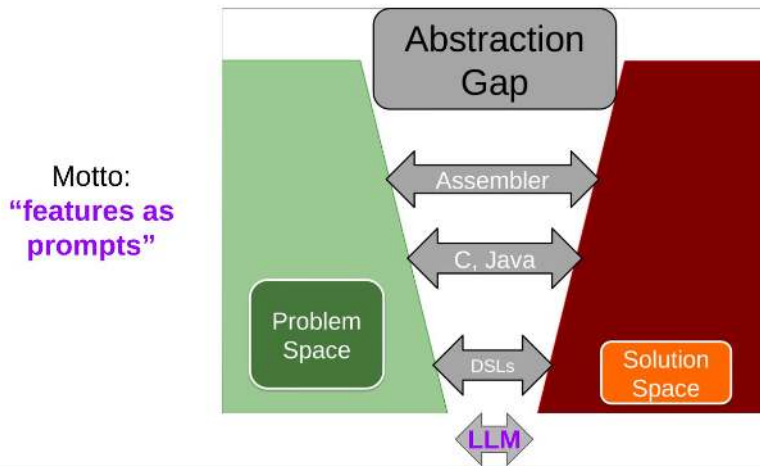
compiler	flags	type	result
gcc		original	z (using long double) = 0x8.0000000000004p-3
gcc		double	z (using double) = 0x1.0000000000001p+0
gcc		float	z (using float) = 0x1p+0
gcc	-mfpmath=387	original	z (using long double) = 0x8.0000000000004p-3
gcc	-mfpmath=387	double	z (using double) = 0x1p+0
gcc	-mfpmath=387	float	z (using float) = 0x1p+0
gcc	-ffast-math	original	z (using long double) = 0x8.0000000000004p-3
gcc	-ffast-math	double	z (using double) = 0x1.0000000000001p+0
gcc	-ffast-math	float	z (using float) = 0x1p+0
gcc	-fno-signed-zeros	original	z (using long double) = 0x8.0000000000004p-3
gcc	-fno-signed-zeros	double	z (using double) = 0x1.0000000000001p+0
gcc	-fno-signed-zeros	float	z (using float) = 0x1p+0
gcc	-fno-trapping-math	original	z (using long double) = 0x8.0000000000004p-3
gcc	-fno-trapping-math	double	z (using double) = 0x1.0000000000001p+0
gcc	-fno-trapping-math	float	z (using float) = 0x1p+0
gcc	-fno-rounding-math	original	z (using long double) = 0x8.0000000000004p-3
gcc	-fno-rounding-math	double	z (using double) = 0x1.0000000000001p+0
gcc	-fno-rounding-math	float	z (using float) = 0x1p+0
clang		original	z (using long double) = 0x8.0000000000004p-3
clang		double	z (using double) = 0x1.0000000000001p+0

retrieve the result of S. Boldo et al.

Potential

LLMs can assist developers in implementing variability in different programming languages (C, Rust, Java, TikZ, etc.) and mechanisms (conditional compilation, feature toggles, command-line parameters, template, etc.). With "features as prompts", there is hope to raise the level of abstraction, increase automation, and bring more flexibility when synthesizing and exploring software variants.

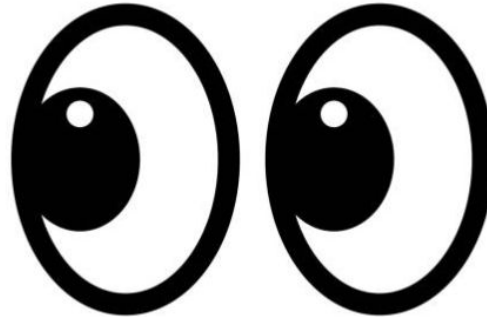
Hypothesis: **Large language models (LLMs)** act as a new variability compiler capable of transforming a high-level specification ("prompt") into variable code, features, generators, configurable systems, etc. written in a given technological space.



Programming without directly manipulating programs, relying almost entirely on AI to manage technical details and create variants of code (customization).

Here is a TikZ code:

```
\tikzset{%
something/.pic={
\tikzset{x=3cm/5,y=3cm/5,shift={(0,-1/3)}}
\useasboundingbox (-1,-1) (1,2);
\fill [BlueGrey900] (0,-2)
.. controls ++(180:3) and ++(0:5/4) .. (-2,0)
arc (270:90:1/5)
.. controls ++(0:2) and ++(180:11/4) .. (0,-2+2/5);
\foreach \i in {-1,1}
\scoped[shift={(1/2*\i,9/4)}, rotate=45*\i]{
\clip [overlay] (0, 5/9) ellipse [radius=8/9];
\clip [overlay] (0,-5/9) ellipse [radius=8/9];
\fill [BlueGrey900] ellipse [radius=1];
\clip [overlay] (0, 7/9) ellipse [radius=10/11];
\clip [overlay] (0,-7/9) ellipse [radius=10/11];
\fill [Purple100] ellipse [radius=1];
};
\fill [BlueGrey900] ellipse [x radius=3/4, y radius=2];
\fill [BlueGrey100] ellipse [x radius=1/3, y radius=1];
\fill [BlueGrey900]
(0,15/8) ellipse [x radius=1, y radius=5/6]
(0, 8/6) ellipse [x radius=1/2, y radius=1/2]
{{shift={(-1/2,-2)}, rotate= 10} ellipse [x radius=1/3, y radius=5/4]}
{{shift={( 1/2,-2)}, rotate=-10} ellipse [x radius=1/3, y radius=5/4]};
\fill [BlueGrey500]
(-1/9,11/8) ellipse [x radius=1/5, y radius=1/5]
(1/9,11/8) ellipse [x radius=1/5, y radius=1/5];
```



Here is a TikZ code:

```
\tikzset{%
something/.pic={
\tikzset{x=3cm/5,y=3cm/5,shift={(0,-1/3)}}
\useasboundingbox (-1,-1) (1,2);
\fill [BlueGrey900] (0,-2)
.. controls ++(180:3) and ++(0:5/4) .. (-2,0)
arc (270:90:1/5)
.. controls ++(0:2) and ++(180:11/4) .. (0,-2+2/5);
\foreach \i in {-1,1}
\scoped[shift={(1/2*\i,9/4)}, rotate=45*\i]{
\clip [overlay] (0, 5/9) ellipse [radius=8/9];
\clip [overlay] (0,-5/9) ellipse [radius=8/9];
\fill [BlueGrey900] ellipse [radius=1];
\clip [overlay] (0, 7/9) ellipse [radius=10/11];
\clip [overlay] (0,-7/9) ellipse [radius=10/11];
\fill [Purple100] ellipse [radius=1];
};
\fill [BlueGrey900] ellipse [x radius=3/4, y radius=2];
\fill [BlueGrey100] ellipse [x radius=1/3, y radius=1];
\fill [BlueGrey900]
(0,15/8) ellipse [x radius=1, y radius=5/6]
(0, 8/6) ellipse [x radius=1/2, y radius=1/2]
{[shift={(-1/2,-2)}, rotate= 10] ellipse [x radius=1/3, y radius=5/4]}
{[shift={( 1/2,-2)}, rotate=-10] ellipse [x radius=1/3, y radius=5/4]};
\fill [BlueGrey500]
(-1/9,11/8) ellipse [x radius=1/5, y radius=1/5]
```

**SE task = software
customization/configuration | program
understanding**


```

Here is a TikZ code:
\tikzset{%
cat/.pic={
\tikzset{x=3cm/5,y=3cm/5,shift={{(0,-1/3)}}
\useasboundingbox (-1,-1) (1,2);
\fill [BlueGrey900] (0,-2)
.. controls ++(180:3) and ++(0:5/4) .. (-2,0)
arc (270:90:1/5)
.. controls ++(0:2) and ++(180:11/4) .. (0,-2+2/5);
\foreach \i in {-1,1}
\scoped[shift={{(1/2*\i,9/4)}, rotate=45*\i}{
\clip [overlay] (0, 5/9) ellipse [radius=8/9];
\clip [overlay] (0,-5/9) ellipse [radius=8/9];
\fill [BlueGrey900] ellipse [radius=1];
\clip [overlay] (0, 7/9) ellipse [radius=10/11];
\clip [overlay] (0,-7/9) ellipse [radius=10/11];
\fill [Purple100] ellipse [radius=1];
};
\fill [BlueGrey900] ellipse [x radius=3/4, y radius=2];
\fill [BlueGrey100] ellipse [x radius=1/3, y radius=1];
\fill [BlueGrey900]
(0,15/8) ellipse [x radius=1, y radius=5/6]
(0, 8/6) ellipse [x radius=1/2, y radius=1/2]
{{shift={{(-1/2,-2)}, rotate= 10} ellipse [x radius=1/3, y radius=5/4]}
{{shift={{ 1/2,-2)}, rotate=-10} ellipse [x radius=1/3, y radius=5/4]};
\fill [BlueGrey500]
(-1/9,11/8) ellipse [x radius=1/5, y radius=1/5]
( 1/9,11/8) ellipse [x radius=1/5, y radius=1/5];
\fill [Purple100]
(0,12/8) ellipse [x radius=1/10, y radius=1/5]
(0,12/8+1/9) ellipse [x radius=1/5, y radius=1/10];
\foreach \i in {-1,1}
\scoped[shift={{(1/2*\i,2)}, rotate=35*\i}{
\clip [overlay] (0, 1/7) ellipse [radius=2/7];
\clip [overlay] (0,-1/7) ellipse [radius=2/7];
\fill [Yellow50] ellipse [radius=1];
};
\scoped{
\clip (-1,-2) rectangle ++(2,1);
\fill [BlueGrey900] (0,-2) ellipse [radius=1/2];
\fill [Grey100]
(-1/2,-2) ellipse [x radius=1/3, y radius=1/4]
( 1/2,-2) ellipse [x radius=1/3, y radius=1/4];
};
\foreach \i in {-1,1}
\foreach \j in {-1,0,1}
\fill [Grey100, shift={{(0,11/8)}, xscale=\i, rotate=\j*15,
shift=(0:1/2)\i

```

SE task = software customization/configuration | program understanding



Here is a TikZ code:

```
\tikzset{%
  cat/.pic={
    \tikzset{x=3cm/5,y=3cm/5,shift={{(0,-1/3)}}}
    \useasboundingbox (-1,-1) (1,2);
    \fill [BlueGrey900] (0,-2)
      .. controls ++(180:3) and ++(0:5/4) .. (-2,0)
      arc (270:90:1/5)
      .. controls ++(0:2) and ++(180:11/4) .. (0,-2+2/5);
    \foreach \i in {-1,1}
      \scoped[shift={{(1/2*\i,9/4)}}, rotate=45*\i]{
        \clip [overlay] (0, 5/9) ellipse [radius=8/9];
        \clip [overlay] (0,-5/9) ellipse [radius=8/9];
        \fill [BlueGrey900] ellipse [radius=1];
        \clip [overlay] (0, 7/9) ellipse [radius=10/11];
        \clip [overlay] (0,-7/9) ellipse [radius=10/11];
        \fill [Purple100] ellipse [radius=1];
      };
    \fill [BlueGrey900] ellipse [x radius=3/4, y radius=2];
    \fill [BlueGrey100] ellipse [x radius=1/3, y radius=1];
    \fill [BlueGrey900]
      (0,15/8) ellipse [x radius=1, y radius=5/6]
      (0, 8/6) ellipse [x radius=1/2, y radius=1/2]
      {{shift={{(-1/2,-2)}}, rotate= 10} ellipse [x radius=1/3, y radius=5/6]}
      {{shift={{( 1/2,-2)}}, rotate=-10} ellipse [x radius=1/3, y radius=5/6]};
    \fill [BlueGrey500]
      (-1/9,11/8) ellipse [x radius=1/5, y radius=1/5]
      ( 1/9,11/8) ellipse [x radius=1/5, y radius=1/5];
    \fill [Purple100]
      (0,12/8) ellipse [x radius=1/10, y radius=1/5]
      (0,12/8+1/9) ellipse [x radius=1/5, y radius=1/10];
    \foreach \i in {-1,1}
      \scoped[shift={{(1/2*\i,2)}}, rotate=35*\i]{
        \clip [overlay] (0, 1/7) ellipse [radius=2/7];
        \clip [overlay] (0,-1/7) ellipse [radius=2/7];
        \fill [Yellow50] ellipse [radius=1];
      };
  };
}
```

I'd like to change the color of the cat's tail



cat

SE task = software customization/configuration | program understanding

Here is a TikZ code:

```
\tikzset{%
cat/.pic={
\tikzset{x=3cm/5,y=3cm/5,shift={{(0,-1/3)}}}
\useasboundingbox (-1,-1) (1,2);
\fill [BlueGrey900] (0,-2)
.. controls ++(180:3) and ++(0:5/4) .. (-2,0)
arc (270:90:1/5)
.. controls ++(0:2) and ++(180:11/4) .. (0,-2+2/5);
\foreach \i in {-1,1}
\scoped[shift={{(1/2*\i,9/4)}}, rotate=45*\i]{
\clip [overlay] (0, 5/9) ellipse [radius=8/9];
\clip [overlay] (0,-5/9) ellipse [radius=8/9];
\fill [BlueGrey900] ellipse [radius=1];
\clip [overlay] (0, 7/9) ellipse [radius=10/11];
\clip [overlay] (0,-7/9) ellipse [radius=10/11];
\fill [Purple100] ellipse [radius=1];
};
\fill [BlueGrey900] ellipse [x radius=3/4, y radius=2];
\fill [BlueGrey100] ellipse [x radius=1/3, y radius=1];
\fill [BlueGrey900]
(0,15/8) ellipse [x radius=1, y radius=5/6]
(0, 8/6) ellipse [x radius=1/2, y radius=1/2]
{{shift={{(-1/2,-2)}}, rotate= 10} ellipse [x radius=1/3, y radius=5/4]}
{{shift={{( 1/2,-2)}}, rotate=-10} ellipse [x radius=1/3, y radius=5/4]};
\fill [BlueGrey500]
(-1/9,11/8) ellipse [x radius=1/5, y radius=1/5]
( 1/9,11/8) ellipse [x radius=1/5, y radius=1/5];
\fill [Purple100]
(0,12/8) ellipse [x radius=1/10, y radius=1/5]
(0,12/8+1/9) ellipse [x radius=1/5, y radius=1/10];
\foreach \i in {-1,1}
\scoped[shift={{(1/2*\i,2)}}, rotate=35*\i]{
\clip [overlay] (0, 1/7) ellipse [radius=2/7];
\clip [overlay] (0,-1/7) ellipse [radius=2/7];
\fill [Yellow50] ellipse [radius=1];
};
};
```

I'd like to change the color of the cat's tail

(LLM) To change the color of the tail, you need to modify the fill color specified in the first 'fill' command, which draws the tail shape. In the provided code, the tail color is set to 'BlueGrey900'. You can replace this color with any other color you'd like. Here's an example with the tail color changed to 'red':

```
\tikzset{%
cat/.pic={
\tikzset{x=3cm/5,y=3cm/5,shift={{(0,-1/3)}}}
\useasboundingbox (-1,-1) (1,2);
\fill [red] (0,-2) % Change the color here
.. controls ++(180:3) and ++(0:5/4) .. (-2,0)
arc (270:90:1/5)
.. controls ++(0:2) and ++(180:11/4) .. (0,-2+2/5);
% The rest of the code remains the same
```

added by
ChatGPT!



change as much as possible elements, applying different styles, colors, and transformations. Give me the new TikZ code and explain me the changes

Can you write a Python program that would apply the 9 proposed changes? I recommend using a templating engine in such a way the 9 changes can be applied and configured

```
cat/.pic={
  \tikzset{x=3cm/5,y=3cm/5,shift={(0,-1/3)}}
  \useasboundingbox (-1,-1) (1,2);
  \fill [<< tail_color >>] (0,-2)
    .. controls ++(180:3) and ++(0:5/4) .. (-2,0)
    arc (270:90:1/5)
    .. controls ++(0:2) and ++(180:11/4) .. (0,-2+2/5);
  ...
  \fill [<< central_body_color >>] ellipse [x radius
    =3/4, y radius=2];
  \fill [<< inner_body_color >>] ellipse [x radius=1/3,
    y radius=1];
  \fill [<< head_paws_color >>]
  ...
}
```

```
import random

def random_rgb_color():
    return f"rgb({random.randint(0, 255)},{random.randint(0, 255)},{random.randint(0, 255)})"

def random_scale(min_value=0.5, max_value=2.0):
    return random.uniform(min_value, max_value)

output = template.render(
    main_body_color=random_rgb_color(),
    head_stripe_color=random_rgb_color(),
    inner_stripe_color=random_rgb_color(),
    head_stripe_scale=random_scale(),
    head_surround_color=random_rgb_color(),
    head_inner_color=random_rgb_color(),
    nose_main_color=random_rgb_color(),
    eyes_color=random_rgb_color(),
    nose_color=random_rgb_color(),
    whisker_spots_color=random_rgb_color(),
    whisker_spots_scale=random_scale(),
    face_lower_part_color=random_rgb_color(),
    face_inner_ellipses_color=random_rgb_color(),
    whiskers_color=random_rgb_color(),
)

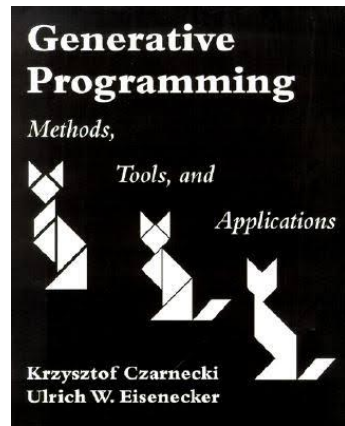
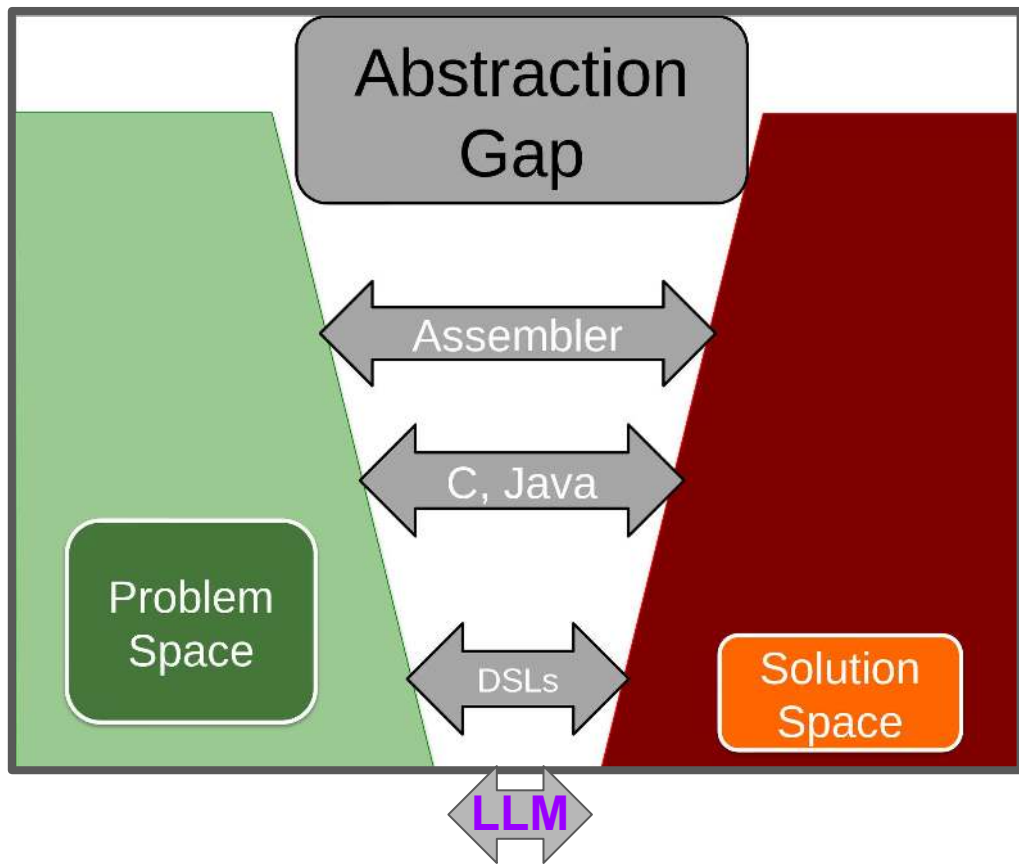
print(output)
```



Hypothesis: **Large language models (LLMs)** act as a new variability compiler capable of transforming a high-level specification (“prompt”) into variable code, features, generators, configurable systems, etc. written in a given technological space.

Motto:
“features as prompts”

Mathieu Acher, Jean-Marc Jézéquel, and José A. Galindo. 2023. On Programming Variability with Large Language Model-based Assistant. In SPLC 2023



LLMs can dramatically fail at times, making mistakes that human developers need to carefully verify and correct.





cat

I'd like to change the color of the cat's tail in red

Here is a TikZ code:

```
\tikzset{%
cat/.pic={
\tikzset{x=3cm/5,y=3cm/5,shift={(0,-1/3)}}
\useasboundingbox (-1,-1) (1,2);
\fill [BlueGrey900] (0,-2)
.. controls ++(180:3) and ++(0:5/4) .. (-2,0)
arc (270:90:1/5)
.. controls ++(0:2) and ++(180:11/4) .. (0,-2+2/5);
\foreach \i in {-1,1}
\scoped[shift={(1/2*\i,9/4)}, rotate=45*\i]{
\clip [overlay] (0,5/9) ellipse [radius=8/9];
\clip [overlay] (0,-5/9) ellipse [radius=8/9];
\fill [BlueGrey900] ellipse [radius=1];
\clip [overlay] (0,7/9) ellipse [radius=10/11];
\clip [overlay] (0,-7/9) ellipse [radius=10/11];
\fill [Purple100] ellipse [radius=1];
};
\fill [BlueGrey900] ellipse [x radius=3/4, y radius=2];
\fill [BlueGrey100] ellipse [x radius=1/3, y radius=1];
\fill [BlueGrey900]
(0,15/8) ellipse [x radius=1, y radius=5/6]
(0,8/6) ellipse [x radius=1/2, y radius=1/2]
{shift={(-1/2,-2)}, rotate=10} ellipse [x radius=1/3, y radius=5/4];
{shift={(1/2,-2)}, rotate=-10} ellipse [x radius=1/3, y radius=5/4];
\fill [BlueGrey500]
(-1/9,11/8) ellipse [x radius=1/5, y radius=1/5]
(1/9,11/8) ellipse [x radius=1/5, y radius=1/5];
\fill [Purple100]
(0,12/8) ellipse [x radius=1/10, y radius=1/5]
(0,12/8+1/9) ellipse [x radius=1/5, y radius=1/10];
\foreach \i in {-1,1}
\scoped[shift={(1/2*\i,2)}, rotate=35*\i]{
\clip [overlay] (0,1/7) ellipse [radius=2/7];
\clip [overlay] (0,-1/7) ellipse [radius=2/7];
\fill [Yellow50] ellipse [radius=1];
};
\scoped{
\clip (-1,-2) rectangle ++(2,1);
\fill [BlueGrey900] (0,-2) ellipse [radius=1/2];
\fill [Grey100]
(-1/2,-2) ellipse [x radius=1/3, y radius=1/4]
(1/2,-2) ellipse [x radius=1/3, y radius=1/4];
};
\foreach \i in {-1,1}
\foreach \j in {-1,0,1}
\fill [Grey100, shift={(0,11/8)}, xscale=\i, rotate=\j*15,
shift=(0:1/2)]
ellipse [x radius=1/3, y radius=1/64];
};
}
```



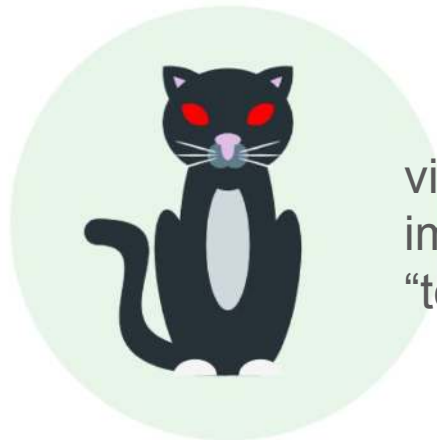
cat



cat

I'd like to change the color of the cat's tail in red

MA You



cat

visual abilities to implement a "testing oracle"

what is the color of the cat's tail? is it red?

MA ChatGPT

The cat's tail is black, not red. The only red part on the cat is its nose.

MA You

what are eyes colors?

MA ChatGPT

The cat's eyes are red.

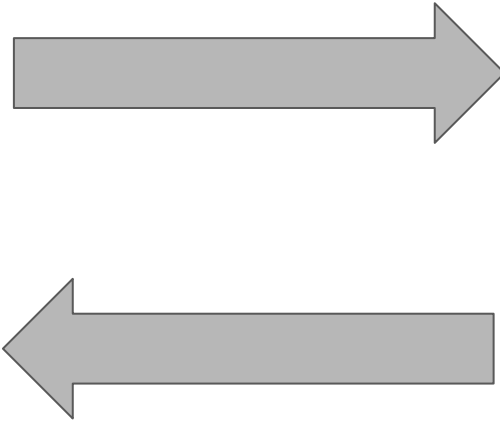


Here is a TikZ code:

```

\tikzset{%
cat/.pic={
\tikzset{x=3cm/5,y=3cm/5,shift={(0,-1/3)}}
\useasboundingbox (-1,-1) (1,2);
\fill [BlueGrey900] (0,-2)
.. controls ++(180:3) and ++(0:5/4) .. (-2,0)
arc (270:90:1/5)
.. controls ++(0:2) and ++(180:11/4) .. (0,-2+2/5);
\foreach \i in {-1,1}
\scoped[shift={(1/2*\i,9/4)}, rotate=45*\i]{
\clip [overlay] (0, 5/9) ellipse [radius=8/9];
\clip [overlay] (0,-5/9) ellipse [radius=8/9];
\fill [BlueGrey900] ellipse [radius=1];
\clip [overlay] (0, 7/9) ellipse [radius=10/11];
\clip [overlay] (0,-7/9) ellipse [radius=10/11];
\fill [Purple100] ellipse [radius=1];
};
\fill [BlueGrey900] ellipse [x radius=3/4, y radius=2];
\fill [BlueGrey100] ellipse [x radius=1/3, y radius=1];
\fill [BlueGrey900]
(0,15/8) ellipse [x radius=1, y radius=5/6]
(0, 8/6) ellipse [x radius=1/2, y radius=1/2]
{shift={(-1/2,-2)}, rotate= 10} ellipse [x radius=1/3, y radius=5/4];
{shift={( 1/2,-2)}, rotate=-10} ellipse [x radius=1/3, y radius=5/4];
\fill [BlueGrey500]
(-1/9,11/8) ellipse [x radius=1/5, y radius=1/5]
( 1/9,11/8) ellipse [x radius=1/5, y radius=1/5];
\fill [Purple100]
(0,12/8) ellipse [x radius=1/10, y radius=1/5]
(0,12/8+1/9) ellipse [x radius=1/5, y radius=1/10];
\foreach \i in {-1,1}
\scoped[shift={(1/2*\i,2)}, rotate=35*\i]{
\clip [overlay] (0, 1/7) ellipse [radius=2/7];
\clip [overlay] (0,-1/7) ellipse [radius=2/7];
\fill [Yellow50] ellipse [radius=1];
};
\scoped{
\clip (-1,-2) rectangle ++(2,1);
\fill [BlueGrey900] (0,-2) ellipse [radius=1/2];
\fill [Grey100]
(-1/2,-2) ellipse [x radius=1/3, y radius=1/4]
( 1/2,-2) ellipse [x radius=1/3, y radius=1/4];
};
\foreach \i in {-1,1}
\foreach \j in {-1,0,1}
\fill [Grey100, shift={(0,11/8)}, xscale=\i, rotate=\j*15,
shift=(0:1/2)]
ellipse [x radius=1/3, y radius=1/64];
}

```





cat

```

Here is a TikZ code:
\tikzset{%
cat/.pic={
\tikzset{x=3cm/5,y=3cm/5,shift={{(0,-1/3)}}
\useasboundingbox (-1,-1) (1,2);
\fill [BlueGrey900] (0,-2)
.. controls ++(180:3) and ++(0:5/4) .. (-2,0)
arc (270:90:1/5)
.. controls ++(0:2) and ++(180:11/4) .. (0,-2+2/5);
\foreach \i in {-1,1}
\scoped[shift={{(1/2*\i,9/4)}, rotate=45*\i}{
\clip [overlay] (0, 5/9) ellipse [radius=8/9];
\clip [overlay] (0,-5/9) ellipse [radius=8/9];
\fill [BlueGrey900] ellipse [radius=1];
\clip [overlay] (0, 7/9) ellipse [radius=10/11];
\clip [overlay] (0,-7/9) ellipse [radius=10/11];
\fill [Purple100] ellipse [radius=1];
};
\fill [BlueGrey900] ellipse [x radius=3/4, y radius=2];
\fill [BlueGrey100] ellipse [x radius=1/3, y radius=1];
\fill [BlueGrey900]
(0,15/8) ellipse [x radius=1, y radius=5/6]
(0, 8/6) ellipse [x radius=1/2, y radius=1/2]
{shift={{(-1/2,-2)}, rotate=10} ellipse [x radius=1/3, y radius=5/4];
{shift={{(1/2,-2)}, rotate=-10} ellipse [x radius=1/3, y radius=5/4];
\fill [BlueGrey500]
(-1/9,11/8) ellipse [x radius=1/5, y radius=1/5]
(1/9,11/8) ellipse [x radius=1/5, y radius=1/5];
\fill [Purple100]
(0,12/8) ellipse [x radius=1/10, y radius=1/5]
(0,12/8+1/9) ellipse [x radius=1/5, y radius=1/10];
\foreach \i in {-1,1}
\scoped[shift={{(1/2*\i,2)}, rotate=35*\i}{
\clip [overlay] (0, 1/7) ellipse [radius=2/7];
\clip [overlay] (0,-1/7) ellipse [radius=2/7];
\fill [Yellow50] ellipse [radius=1];
};
\scoped{
\clip (-1,-2) rectangle ++(2,1);
\fill [BlueGrey900] (0,-2) ellipse [radius=1/2];
\fill [Grey100]
(-1/2,-2) ellipse [x radius=1/3, y radius=1/4]
(1/2,-2) ellipse [x radius=1/3, y radius=1/4];
};
\foreach \i in {-1,1}
\foreach \j in {-1,0,1}
\fill [Grey100, shift={{(0,11/8)}}, xscale=\i, rotate=\j*15,
shift=(0:1/2)]
ellipse [x radius=1/3, y radius=1/64];
};
}

```



“I’d like to have a better red, more dark...” (kids)

Customization

Ears, tail, and whisker

tail :	<input type="text" value="red"/>
outer ear :	<input type="text" value="BlueGrey900"/>
inner ear :	<input type="text" value="Purple100"/>
whiskers :	<input type="text" value="Grey100"/>

Beautify the cat!



cat

<https://www.youtube.com/watch?v=kSdLSqwZK7Q>

(plenty of experiments with LLMs, programming languages, students... I'll blog about it soon!)



My banker suggested this investment:

you give me $e \approx 2.71828...$ €,

the following year, I take €1 in fees and multiply by 1,

the following year, I take 1€ in fees and multiply by 2,

the following year, I take a €1 charge and multiply by 3,

...

after n years, I take a €1 fee and multiply by n .

To get my money back, there's a €1 fee.

In 50 years, when I retire, how much money will I have?

Synthesize and execute a Python code to answer the question



After 50 years, your investment would be approximately -5.56×10^{58} euros, meaning the investment results in an astronomically large negative balance. This investment strategy leads to catastrophic losses. [-]



MA You
I have the following Python code:

```
import folium
```

```
# Given coordinates for the cities (these are example coordinates and may not be accurate)
```

```
# Normally we would obtain these through a geocoding service  
coordinates = {
```

```
    "Rennes": [48.1173, -1.6778],  
    "Paris": [48.8566, 2.3522],  
    "Saclay": [48.7290, 2.1912],  
    "Lille": [50.6292, 3.0573],  
    "Bordeaux": [44.8378, -0.5792],  
    "Sophia": [43.6222, 7.0500], # This is an approximation for Sophia Antipolis  
    "Nancy": [48.6921, 6.1844]  
}
```

```
# Initialize a map centered around France
```

```
map_france = folium.Map(location=[46.603354, 1.888334], zoom_start=6)
```

```
# Add markers for each team
```

```
add_team_marker(map_france, "DIVERSE", "Rennes")  
add_team_marker(map_france, "PICUBE", "Paris")  
add_team_marker(map_france, "ARGO", "Paris")  
add_team_marker(map_france, "DEDUCTEAM", "Saclay")  
add_team_marker(map_france, "EVEREF", "Lille")  
add_team_marker(map_france, "FLOWERS", "Bordeaux")  
add_team_marker(map_france, "MNEMOSYNE", "Bordeaux")  
add_team_marker(map_france, "SPIRALS", "Lille")  
add_team_marker(map_france, "STAMP", "Sophia")  
add_team_marker(map_france, "SED", "Nancy")
```

```
# Since we cannot display the folium map directly in this environment,  
# we'll save it to an HTML file which you can then open in a web browser.
```

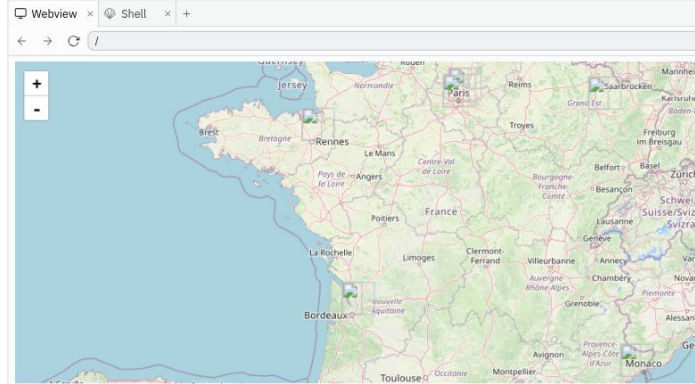
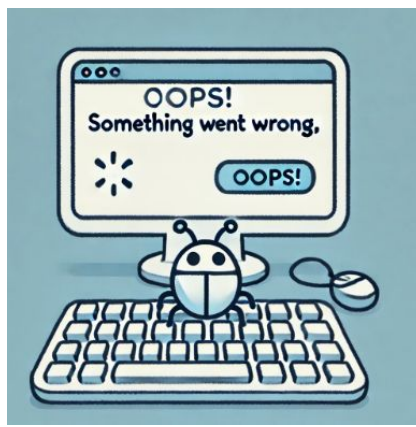
```
map_file_path = '/mnt/data/france_teams_map.html'  
map_france.save(map_file_path)  
map_file_path
```

Please write a JavaScript program, using Leaflet.js

CP ChatGPT

Creating a JavaScript program using Leaflet.js to achieve a similar functionality as your Python code with Folium is a good way to visualize geographic data on web platforms. Leaflet is a leading open-source JavaScript library for mobile-friendly interactive maps.

Here's a basic example of how you could write such a program:



SE task: modernization/migration

```
index.html x scripts.js x  
script.js x f addCityMarker x popupText  
1 var map = L.map('map').setView([46.603354, 1.888334], 6);  
2  
3 // Add OpenStreetMap tiles:  
4 L.tileLayer('https://{s}.tile.openstreetmap.org/{z}/{x}/{y}.png', {  
5   maxZoom: 19,  
6   attribution: '© OpenStreetMap contributors'  
7 }).addTo(map);  
8  
9 // Teams and their respective cities  
10 var teams = [  
11   { city: "Rennes", team: "DIVERSE" },  
12   { city: "Paris", team: "PICUBE" },  
13   { city: "Paris", team: "ARGO" },  
14   { city: "Saclay", team: "DEDUCTEAM" },  
15   { city: "Lille", team: "EVEREF" },  
16   { city: "Bordeaux", team: "FLOWERS" },  
17   { city: "Bordeaux", team: "MNEMOSYNE" },  
18   { city: "Lille", team: "SPIRALS" },  
19  
20 ]  
21  
22 <!DOCTYPE html>  
23 <html>  
24 <head>  
25   <title>Team Locations in France</title>  
26   <link rel="stylesheet" href="https://unpkg.com/leaflet@1.0.0/dist/leaflet.css" />  
27   <script src="https://unpkg.com/leaflet@1.0.0/dist/leaflet.js"></script>  
28   <style>  
29     #map {  
30       height: 400px;  
31     }  
32 </style>  
33 </head>  
34 <div id="map"></div>  
35 <script src="script.js"></script>  
36 </body>  
37 </html>  
38  
39  
40 // Helper function to get coordinates of a city  
41 function getCoordinates(city) {  
42   var coordinates = {  
43     "Rennes": [48.1173, -1.6778],  
44     "Paris": [48.8566, 2.3522],  
45     "Saclay": [48.7290, 2.1912],  
46     "Lille": [50.6292, 3.0573],  
47     "Bordeaux": [44.8378, -0.5792],  
48     "Sophia": [43.6222, 7.0500],  
49     "Nancy": [48.6921, 6.1844]  
50   };  
51   return coordinates[city];  
52  
53 // Add markers for each city  
54 object.keys(teamsByCity).forEach(function(city) {  
55   addCityMarker(city, teamsByCity[city]);  
56 });  
57  
58
```

<https://arxiv.org/abs/2210.14699>

Piloting Copilot and Codex: Hot Temperature, Cold Prompts, or Black Magic?

```
def choose_num(x, y):  
    """  
    This function takes two positive  
    numbers x and y and returns the  
    biggest even integer number that  
    is in the range [x, y] inclusive.  
    If there's no such number, then  
    the function should return -1.  
  
    For example:  
    choose_num(12, 15) = 14  
    choose_num(13, 12) = -1  
    """  
    if x % 2 != 0:  
        x += 1  
    if y % 2 != 0:  
        y -= 1  
    if x > y:  
        return -1  
    if x % 2 == 0:  
        x += 1  
    while x <= y:  
        if x % 2 == 0:  
            return x  
        x += 2  
    return -1
```

(a) Original prompt

```
def choose_num(x, y):  
    """  
    This function takes two positive  
    numbers x and y and returns the  
    biggest even integer number that  
    is in the range [x, y] inclusive.  
    If there's no such number, then  
    the function should return -1.  
  
    For example:  
    choose_num(12, 15) = 14  
    choose_num(13, 12) = -1  
  
    Write a quick algorithm to  
    solve this problem.  
    """  
    if x % 2 == 0:  
        x = x  
    else:  
        x = x + 1  
    if y % 2 == 0:  
        y = y  
    else:  
        y = y - 1  
    if x > y:  
        return -1  
    else:  
        return max(x, y)
```

(b) Modified prompt 1

```
def choose_num(x, y):  
    """  
    This function takes two positive  
    numbers x and y and returns the  
    biggest even integer number that  
    is in the range [x, y] inclusive.  
    If there's no such number, then  
    the function should return -1.  
    """  
    if x % 2 == 1:  
        x += 1  
    if y % 2 == 0:  
        y -= 1  
    for i in range(x, y + 1, 2):  
        if x <= i <= y:  
            return i  
    return -1
```

(c) Modified prompt 2

Jean-Baptiste Döderlein
ENS Rennes
Rennes, France
jean-baptiste.doderlein@ens-rennes.fr

Djamel Eddine Khelladi
CNRS, Univ Rennes, IRISA, Inria
Rennes, France
djamel-eddine.khelladi@irisa.fr

Mathieu Acher
Univ Rennes, IUF, CNRS, Inria, IRISA
Rennes, France
mathieu.acher@irisa.fr

Benoit Combemale
Univ Rennes, CNRS, Inria, IRISA
Rennes, France
benoit.combemale@irisa.fr



RQ₃ For both HumanEval and LeetCode, varying the temperature can significantly improve Codex performance, up to the point Codex with a configured temperature (not necessarily the best one) can strongly outperform Copilot. The optimal temperature depends on the chosen prompt variation, the number of generated codes, and to a lesser extent the targeted programming languages.

RQ₄ A proper configuration of the temperature and prompt variation can significantly improve the results of Codex. A difficulty, though, is that the best configurations differ for a given pass@k. A general recommendation is to use low temperatures (0.0, 0.2) for pass@1, and higher temperatures (0.6 to 1.0) for pass@10 and @100.

RQ₅ Tuning temperature and prompt variation per problem can provide strong performance improvements: up to 79.27% at pass@1 compared to 22.44% for the default configuration and outperforming best average configuration of RQ4 (52.67%).

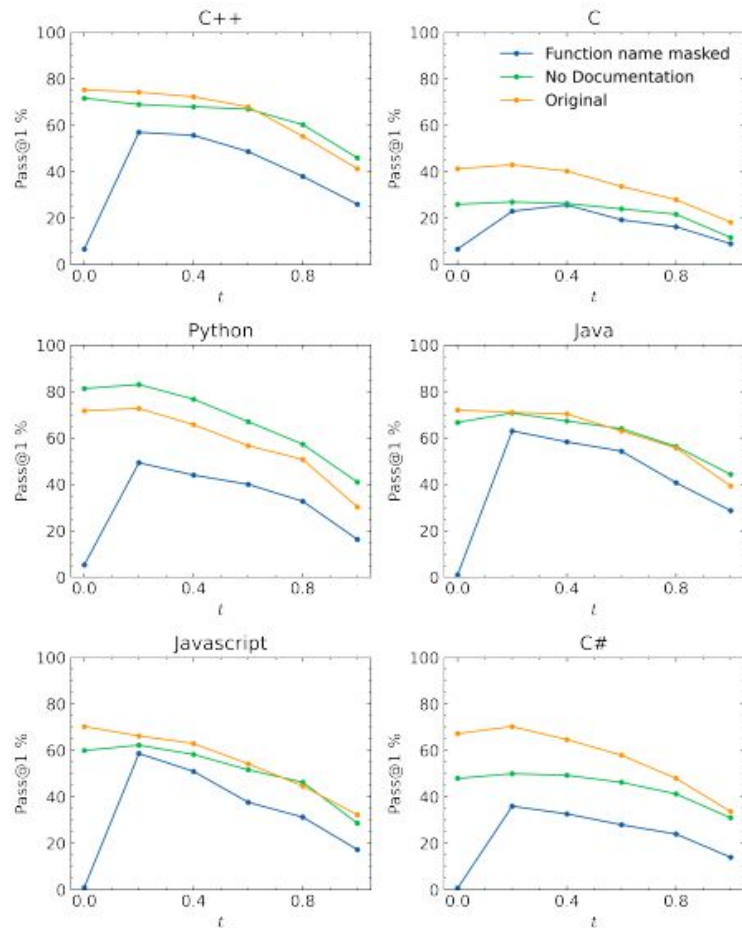


Fig. 3: Pass@1 Evaluation of Codex on Leetcode dataset; $temp \in [0.0, 0.2, 0.4, 0.6, 0.8, 1]$

Challenges

- LLMs are:
 - brittle, leading to faulty, under-optimized, or unreliable code (including tests!)
 - under-trained on “small” languages or not up-to-date with (knowledge of) new libraries or software versions
 - LLMs lack planning and reasoning capabilities
- Challenges of improving LLMs:
 - Better “language” than prompt?
 - more verification/validation tools to verify the outcome?
 - fine-tuning?
- Benchmarking (SE task specific, no data contamination, etc.)
- Open source LLMs (CodeCommons with SoftwareHeritage)



LLM4Code: Reliable and Productive Code Assistants
Based on Large Language Models

Project Leads

DIVERSE	Rennes	Mathieu Acher	mathieu.acher@inria.fr
PICUBE	Paris	Guillaume Baudart	guillaume.baudart@inria.fr

Internal Partners (EP/CRI)

ARGO	Paris	Marc Lelarge	marc.lelarge@inria.fr
DEDUCTEAM	Saclay	Anthony Bordg	anthony.bordg@inria.fr
EVREF	Lille	Anne Etien	anne.etien@univ-lille.fr
FLOWERS	Bordeaux	Pierre-Yves Oudeyer	pierre-yves.oudeyer@inria.fr
MNEMOSYNE	Bordeaux	Xavier Hinaut	xavier.hinaut@inria.fr
SPIRALS	Lille	Clément Quinton	clement.quinton@inria.fr
STAMP	Lille	Cyril Cohen	cyril.cohen@inria.fr
SED	Nancy	Laurent Pierron	laurent.pierron@inria.fr

External Partners

SOFTWARE HERITAGE	Roberto di Cosmo	roberto@dicosmo.org
LABRI	Nathanaël Fijalkow	nathanael.fijalkow@labri.fr
	Romain Robbes	romain.robbes@cnrs.fr
SOPRA STERIA	Yann Gloriant	yann.gloriau@soprasteria.com

CODE COMMONS: "LE" BIEN COMMUN FOR SOURCE CODE IN FRANCE!

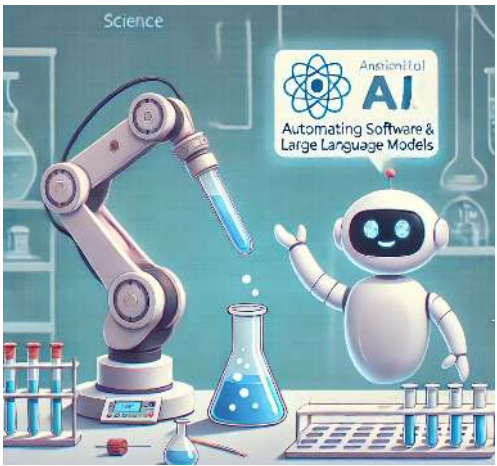
For new generation AI!

starcoder2 Public

(illustration of LLM prompt sensitivity in chess)



How LLMs can be used both to generate variants (e.g., prompts, hypotheses, ideas) and to accelerate scientific discovery through software-based exploration and analysis.



What
is your move?



What is your prompt?

[Event "FIDE World
Championship Match 2024"]
[Site "Los Angeles, USA"]
[Date "2024.12.01"]
[Round "5"]
[White "Kramnik, Vladimir"]
[Black "Nepomniachtchi,
Ian"]
[Result "1-0"]
[WhiteElo "2900"]
[BlackElo "2900"]
[TimeControl
"40/7200:20/3600:900+30"]
[UTCDate "2024.11.27"]
[UTCTime "09:01:25"]
[Variant "Standard"]



[Event "FIDE World
Championship Match 2024"]
[Site "Los Angeles, USA"]
[Date "2024.12.01"]
[Round "5"]
[White "Louapre, David"]
[Black "Giraud, Thibaut"]
[Result "0-1"]
[WhiteElo "1400"]
[BlackElo "1400"]
[TimeControl
"40/7200:20/3600:900+30"]
[UTCDate "2024.11.27"]
[UTCTime "09:01:25"]
[Variant "Standard"]



[Event
Champ
[Site "L
[Date "
[Round
[White
[Black
lan"]
[Resul
[White
[Black
[TimeC
"40/72
[UTCD
[UTCT
[Variar

[Event '
Champ
[Site "L
[Date "
[Round
[White
[Black
[Resul
[White]
[Black]
[TimeC
"40/72(
[UTCD
[UTCTI
[Variar

MrPhi
@MonsieurPhi

Petite note: j'ai remarqué que jusqu'à 1800 Elo, sa sensibilité au prompt est un peu bizarre. Par exemple avec "1-0" (Blanc gagne) il joue le coup correct g6. Mais avec "0-1" (Noir gagne) ou "1/2-1/2", il joue Nf6 (ce qui est illogique vu que Nf6 fait justement perdre les Noirs).

[Translate post](#)

Playground Complete

<pre>[Event "Chess Tournament"] [Site "?"] [Date "2024.01.15"] [Round "1"] [White "Louapre, David"] [Black "Carlsen, Magnus"] [Result "1-0"] [WhiteElo "1800"] [BlackElo "1800"] 1.e4 e5 2.Bc4 Nc6 3.Qh5 g6</pre>	<pre>[Event "Chess Tournament"] [Site "?"] [Date "2024.01.15"] [Round "1"] [White "Louapre, David"] [Black "Carlsen, Magnus"] [Result "0-1"] [WhiteElo "1800"] [BlackElo "1800"] 1.e4 e5 2.Bc4 Nc6 3.Qh5 Nf6</pre>
--	---

11:55 AM · Apr 19, 2024 · 447 Views

1 7

Post your reply Reply

MrPhi @MonsieurPhi · Apr 19
Mais il n'a plus ce comportement bizarre quand le Elo dépasse 2000. Enfin bref ! Il faudrait faire d'autres tests pour déterminer à quel point cette information sur le niveau Elo influence son niveau de jeu, mais c'est déjà intéressant comme petite expérience.

2 9 441

ur prompt?



Hypothesis and Knowledge

Subject of interest/Population

Experiments (variants)

Data analysis

Interpretation/Conclusion



```

# Optional GPT calls
black_title = "[BlackTitle %04d]" if include_black_title else ""
white_title = "[WhiteTitle %04d]" if include_white_title else ""

# Construct the PROM header with the configurable options:
prom_headers = """[Event "FIDE World Championship Match 2021"]
[Site "Los Angeles, USA"]
[Date "2024.12.01"]
[Round "%d"]
[White "%s"]
[Black "%s"]
[result "%d"]
[WhiteElo "%d"]
[WhiteTitle "%s"]
[BlackElo "%d"]
[BlackTitle "%s"]
"""

```

```

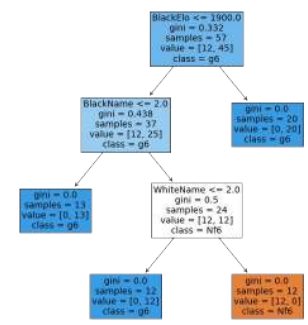
# Define possible values for each parameter
results = ["1-0", "1/2-1/2", "0-1"]
names = ["MagnusCarlsen", "Carlsen", "Carlsen", "Magnus", "Carlsen", "Magnus", "Carlsen", "Magnus"]
elos = [1900, 3400, 1700, 3000, 2900, 3000]
include_title = [True, False]
include_title = [True, False]

```

```

gpt_config = GPTConfig(
    model_gpt="gpt-3.5-turbo-instruct",
    temperature=0.0,
    max_tokens=5,
    chat_gpt=False,
    system_role_message=None # Since it wasn't provided in the original call
)

```



Monsieur Phi is ~right. Sensitivity of LLMs. Do LLMs have “learned” the Elo?

AI-(empowered) scientists: enormous potential to automatically explore large variants spaces (with software and AI) and discover new knowledge or confirm/refute theories.



MrPhi
@MonsieurPhi

Petite note: j'ai remarqué que jusqu'à 1800 Elo, sa sensibilité au prompt est un peu bizarre. Par exemple avec "1-0" (Blanc gagne) il joue le coup correct g6. Mais avec "0-1" (Noir gagne) ou "1/2-1/2", il joue Nf6 (ce qui est illogique vu que Nf6 fait justement perdre les Noirs).

[Translate post](#)

Playground

Complete

Playground

Complete

[Event "Chess Tournament"]

[Site "?"]

[Date "2024.0115"]

[Round "1"]

[White "Louapre, David"]

[Black "Carlsen, Magnus"]

[Result "1-0"]

[WhiteElo "1800"]

[BlackElo "1800"]

1.e4 e5 2.Bc4 Nc6 3.Qh5 g6

[Event "Chess Tournament"]

[Site "?"]

[Date "2024.0115"]

[Round "1"]

[White "Louapre, David"]

[Black "Carlsen, Magnus"]

[Result "0-1"]

[WhiteElo "1800"]

[BlackElo "1800"]

1.e4 e5 2.Bc4 Nc6 3.Qh5 Nf6

?

11:55 AM · Apr 19, 2024 · 447 Views

1



7



Post your reply

Reply



MrPhi @MonsieurPhi · Apr 19

Mais il n'a plus ce comportement bizarre quand le Elo dépasse 2000. Enfin bref ! Il faudrait faire d'autres tests pour déterminer à quel point cette information sur le niveau Elo influence son niveau de jeu, mais c'est déjà intéressant comme petite expérience.

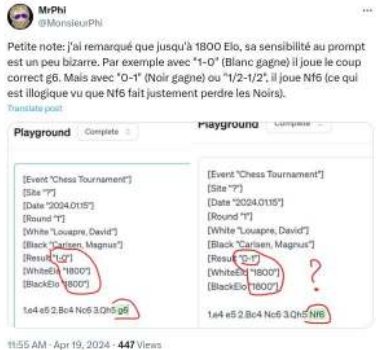
2



9

441





```
# Optional GM titles
black_title = "[BlackTitle \"GM\"]" if include_black_title else ""
white_title = "[WhiteTitle \"GM\"]" if include_white_title else ""

# Construct the PGN header with the configurable options
pgn_headers = f"""[Event "FIDE World Championship Match 2024"]
[Site "Los Angeles, USA"]
[Date "2024.12.01"]
[Round "5"]
[White "{white_name}"]
[Black "{black_name}"]
[Result "{result}"]
[WhiteElo "{white_elo}"]
[BlackElo "{black_elo}"]
[black_title]
[white_title]
[TimeControl "40/7200:20/3600:900+30"]
[UTCDate "2024.11.27"]
[UTCTime "09:01:25"]
[Variant "Standard"]
"""
```

```
# Define possible values for each parameter
results = ["1-0", "1/2-1/2", "0-1"]
# results = ["1-0", "0-1"]
names = ["Nepomniachtchi, Ian", "Kramnik, Vladimir", "Giraud, Thibaut", "Louapre, David", "XXX"]
elos = [1000, 1400, 1700, 1800, 2000, 2900]
# include_title = [True, False]
include_title = [False]
```



```
gpt_config = GPTConfig(
    model_gpt="gpt-3.5-turbo-instruct",
    temperature=0.0,
    max_tokens=5,
    chat_gpt=False,
    system_role_message=None # Since it wasn't provided in the original call
)
```





MrPhi
@MonsieurPhi

Petite note: j'ai remarqué que jusqu'à 1800 Elo, sa sensibilité au prompt est un peu bizarre. Par exemple avec "1-0" (Blanc gagne) il joue le coup correct g6. Mais avec "0-1" (Noir gagne) ou "1/2-1/2", il joue Nf6 (ce qui est illogique vu que Nf6 fait justement perdre les Noirs).

[Translate post](#)

Playground Complete

[Event "Chess Tournament"]
[Site "?"]
[Date "2024.01.15"]
[Round "1"]
[White "Louapre, David"]
[Black "Carlsen, Magnus"]
[Result "1-0"]
[WhiteElo "1800"]
[BlackElo "1800"]

1.e4 e5 2.Bc4 Nc6 3.Qh5 g6

[Event "Chess Tournament"]
[Site "?"]
[Date "2024.01.15"]
[Round "1"]
[White "Louapre, David"]
[Black "Carlsen, Magnus"]
[Result "0-1"]
[WhiteElo "1800"]
[BlackElo "1800"]

1.e4 e5 2.Bc4 Nc6 3.Qh5 Nf6

11:55 AM · Apr 19, 2024 · 447 Views

1 7



Post your reply

Reply

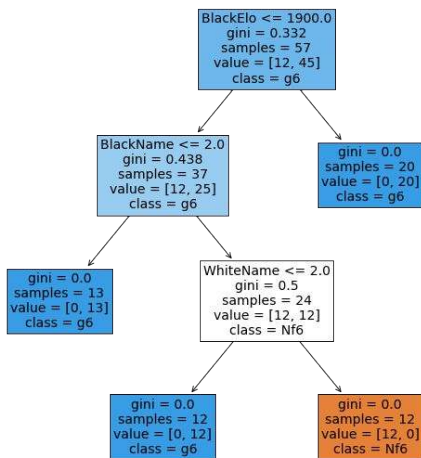


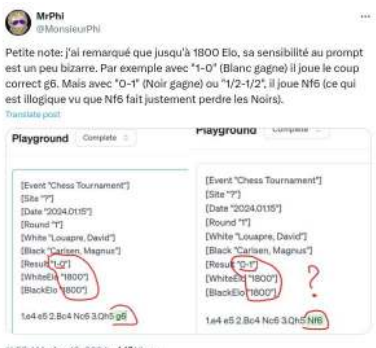
MrPhi @MonsieurPhi · Apr 19

Mais il n'a plus ce comportement bizarre quand le Elo dépasse 2000. Enfin bref ! Il faudrait faire d'autres tests pour déterminer à quel point cette information sur le niveau Elo influence son niveau de jeu, mais c'est déjà intéressant comme petite expérience.

2 441

Result	WhiteName	BlackName	WhiteElo	BlackElo	IncludeWhiteTitle	IncludeBlackTitle	Move
0-1	XXX	Louapre, David	1800	1800	False	False	Nf6
1-0	Louapre, David	XXX	2900	2900	False	False	g6
0-1	Nepomniachtchi, Ian	Louapre, David	2900	2900	False	False	g6
0-1	Nepomniachtchi, Ian	Louapre, David	2000	2000	False	False	g6
0-1	Louapre, David	XXX	1400	1400	False	False	Nf6
1-0	XXX	Nepomniachtchi, Ian	1800	1800	False	False	g6
1-0	Nepomniachtchi, Ian	Louapre, David	2900	2900	False	False	g6
1-0	Louapre, David	Nepomniachtchi, Ian	1800	1800	False	False	g6
0-1	Louapre, David	Nepomniachtchi, Ian	1000	1000	False	False	g6
1-0	Nepomniachtchi, Ian	XXX	2000	2000	False	False	g6
0-1	XXX	Nepomniachtchi, Ian	2000	2000	False	False	g6
0-1	Nepomniachtchi, Ian	XXX	1700	1700	False	False	g6
1-0	Louapre, David	XXX	1700	1700	False	False	Nf6





11:55 AM · Apr 19, 2024 · 447 Views

Post your reply

MPH @MonsieurPH · Apr 19
 Mais il n'a plus ce comportement bizarre quand le Elo dépasse 2000. Enfin bref !! Il faudrait faire d'autres tests pour déterminer à quel point cette information sur le niveau Elo influence son niveau de jeu, mais c'est 0/6/3 intéressant comme petite expérience.

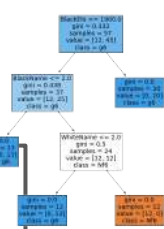
Strategic exploration with modelling and learning (assisted by LLM)

```
onal GM titles
:title = "[BlackTitle \"GM\"]" if include_black_title else ""
:title = "[WhiteTitle \"GM\"]" if include_white_title else ""
```

```
# Construct the PGN header with the configurable options
pgn_headers = f"""[Event "FIDE World Championship Match 2024"]
[Site "Los Angeles, USA"]
[Date "2024.12.01"]
[Round "5"]
[White "{white_name}"]
[Black "{black_name}"]
```

```
gpt_config = GPTConfig
model_gpt="gpt-3.5
temperature=0.0,
max_tokens=5,
chat_gpt=False,
system_role_message=None # Since it wasn't provided in the original call
```

```
# Define possible values for each parameter
results = ["1-0", "1/2-1/2", "0-1"]
# results = ["1-0", "0-1"]
names = ["Nepomniachtchi, Ian", "Kramnik, Vladimir", "Giraud, Thibaut", "Louapre, David", "XXX"]
elos = [1000, 1400, 1700, 1800, 2000, 2900]
# include_title = [True, False]
include_title = [False]
```



Result	WhiteName	BlackName	WhiteElo	BlackElo	IncludeWhiteTitle	IncludeBlackTitle	Move
0-1	XXX	Louapre, David	1800	1800	False	False	Nf6
1-0	Louapre, David	XXX	2900	2900	False	False	g6
0-1	Nepomniachtchi, Ian	Louapre, David	2900	2900	False	False	g6
0-1	Nepomniachtchi, Ian	Louapre, David	2000	2000	False	False	g6
0-1	Louapre, David	XXX	1400	1400	False	False	Nf6
1-0	XXX	Nepomniachtchi, Ian	1800	1800	False	False	g6
1-0	Nepomniachtchi, Ian	Louapre, David	2900	2900	False	False	g6
1-0	Louapre, David	Nepomniachtchi, Ian	1800	1800	False	False	g6
0-1	Louapre, David	Nepomniachtchi, Ian	1000	1000	False	False	g6
1-0	Nepomniachtchi, Ian	XXX	2000	2000	False	False	g6
0-1	XXX	Nepomniachtchi, Ian	2000	2000	False	False	g6
0-1	Nepomniachtchi, Ian	XXX	1700	1700	False	False	g8
1-0	Louapre, David	XXX	1700	1700	False	False	Nf6



Hypothesis and Knowledge

Subject of interest/Population

Experiments (variants)

Data analysis

Interpretation/Conclusion



```

# Optional GPT titles
black_title = "[BlackTitle %04d]" if include_black_title else ""
white_title = "[WhiteTitle %04d]" if include_white_title else ""

# Construct the PGN header with the configurable options
pgn_headers = """[Event "FIDE World Championship Match 2021"]
[Site "Los Angeles, USA"]
[Date "2022.12.01"]
[Round "%d"]
[White "%s"]
[Black "%s"]
[result "%d"]
[WhiteElo "%d"]
[WhiteTitle "%s"]

```

```

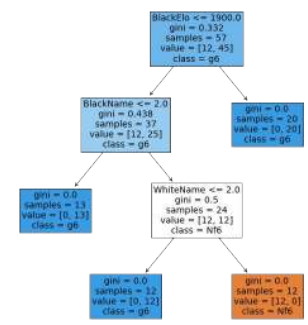
# Define possible values for each parameter
results = ["1-0", "1/2-1/2", "0-1"]
names = ["MagnusCarlsen", "Jan", "Konerik", "Vladimir", "Sviridov", "Theodor", "Kasparov", "David", "XXX"]
elos = [1900, 3400, 1700, 1600, 2900, 2000]
include_title = [True, False]
include_title = [True]

```

```

gpt_config = GPTConfig(
    model_l_gpt="gpt-3.5-turbo-instruct",
    temperature=0.0,
    max_tokens=5,
    chat_gpt=False,
    system_role_message=None # Since it wasn't provided in the original call
)

```



Monsieur Phi is ~right. Sensitivity of LLMs. Do LLMs have “learned” the Elo?

AI-(empowered) scientists:

- this example (chess) is a bit “meta” (we’re assessing AIs)
- still: I find it representative of many scientific studies where we run variants of experiments, analyze data, and communicate insights – AI is useful to assist, but not mandatory here
- many tasks uncovered (eg literature review)

AI scientists can study AIs – no human in the loop?

Hypothesis and Knowledge

Subject of interest/Population

Experiments (variants)

Data analysis

Interpretation/Conclusion

The image shows a social media post from 'MFB' with a French text snippet: "De la robe / et remonte au cas jusqu'à 1800 Fts, sa sensibilité sa grandit...". Below the text is a 'Playground' interface with two columns of code-like parameters: [Event: Chess Tournament], [Site: ?], [Date: 2024.01.07], [Round: ?], [White: Anwar, David], [Black: Carben, Vagner], [Time: 3.7], [WhiteElo: 1207], [BlackElo: 1207], [Type: Bot Not 3 Check]. A red circle highlights the 'BlackElo' field. Below the playground is a video player showing a man speaking.



Scientific knowledge ... ▶

<https://blog.mathieuacher.com/GPTsChessEloRatingLegalMoves/>

AI(-empowered) scientists:
AI scientists can study AIs – no human in the loop?
Automated scientific discovery

Hypothesis and Knowledge

Subject of interest/Population

Experiments (variants)

Data analysis

Interpretation/Conclusion

The image is a composite of two parts. On the left, a person in a light blue shirt is pointing at a whiteboard in a classroom or lecture hall. On the right, a computer screen displays a document titled "Précision intermédiaire - le cas C". The document contains the following text:

Norme ISO 9542.2 (9)

The values of operations with floating operands and values subject to the usual arithmetic conversions and of floating constants are evaluated to a format whose range and precision may be greater than required by the type.

Exemple (Précision étendue)

variable x = 1.0;
variable y = 20.952 + 2e19.64;
variable z = x * y;

peut produire les valeurs suivantes :

- $x = 1 + 2^{-52}$ addition en binary64,
- $z = 1$ addition en binary80 puis stockage en binary64.

Arithmétique à virgule flottante : plate ou terrain de jeu ?
Séverin Belec

Scientific knowledge

AI(-empowered) scientists:
AI scientists can study AIs – no human in the loop?
Automated scientific discovery

As Software is key to Computational Science and Society, numerous challenges and opportunities ahead!

- democratization of software
- reliable/trustable/resilient software
- speeding up productivity (and focusing on tough/creative challenges)
- software to explore variants' space (hypothesis, ideas, products, etc.)
- ...

We need to invest more in software science and software engineering (SE) research

Large Language Models (LLMs) are a transformative technology to raise the level of abstraction, automate SE tasks, and augment/assist software users/developers/scientists → LLM for SE!?

 GitHub Copilot

SE tasks = code synthesis | code completion

<https://arxiv.org/abs/2210.14699>

<https://arxiv.org/abs/2210.14699>

change as much as possible elements applying different styles, colors, and annotations. Give me the new TIKZ code and explain the changes

Can you write a Python program that would apply the 9 proposed changes? I recommend using templating engine in such a way the 9 changes can be applied and configured

```

\set{style}
\usepackage{tikz}
\usetikzlibrary{arrows,calc,decorations,positioning,shadows,shadings}
\begin{tikzpicture}
\draw[fill=red,opacity=0.5] (0,0) circle (1cm);
\draw[fill=blue,opacity=0.5] (1,0) circle (1cm);
\draw[fill=green,opacity=0.5] (2,0) circle (1cm);
\draw[fill=yellow,opacity=0.5] (3,0) circle (1cm);
\draw[fill=orange,opacity=0.5] (4,0) circle (1cm);
\draw[fill=purple,opacity=0.5] (5,0) circle (1cm);
\draw[fill=black,opacity=0.5] (6,0) circle (1cm);
\draw[fill=white,opacity=0.5] (7,0) circle (1cm);
\draw[fill=gray,opacity=0.5] (8,0) circle (1cm);
\end{tikzpicture}

```

SE task = software customization/configuration | program understanding

Piloting Copilot and Codex: Hot Temperature, Cold Prompts, or Black Magic?



Generative AI for Generative Programming: Automating Code Variants and Exploring the Boundaries of LLMs

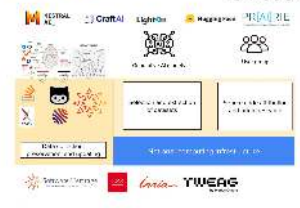
LLM4Code: Reliable and Productive Code Assistants
Based on Large Language Models



Project Leads			
DIVINUS	Reims	Mathieu Acher	nathieu.acher@inria.fr
PARIBR	Paris	Gaëlleme Roudot	gael@lewis.bondard.inria.fr
Internal Partners (EP/CRF)			
ARDO	Paris	Marc Leclercq	marc.leclercq@inria.fr
DEMOICOM	Sudbury	Anthony Borop	anthony.borop@inria.fr
EVREZ	Lille	Amel Eloua	amel.eloua@univ-lille.fr
FLAMINGO	Reims	Pierre-Yves Oudot	pierre-yves.oudot@inria.fr
SYSTEMONE	Reims	Nicolas Hamon	nicolas.hamon@inria.fr
SPIRALIS	Lille	Clement Quinson	clement.quinson@inria.fr
SIAM3	Saphia	Cyril Cohen	cyril.cohen@inria.fr
SED	Nancy	Laurent Pierrot	laurent.pierrot@inria.fr

External Partners		
SOFTWARE HERITAGE	Roberto Di Cosimo	robertodicosimo.org
LADRI	Nathanaël Pijolle	nathanael.pijolle@ladri.fr
SOFTAL	Romain Robbes	romain.robbes@univ-lille.fr
SOFTAL	Yann Eliezer	yann.eliezer@softal.fr

CODE COMMONS: "LE" BIEN COMMUN FOR SOURCE CODE IN FRANCE!




Backup slides

How LLMs can assist in modernizing legacy applications, such as those written in COBOL or in deprecated technologies.

Polyglot?

Python => JavaScript+HTML+CSS

You
I have the following Python code:

```
import folium
```

```
# Given coordinates for the cities (these are example coordinates and may not be accurate)
```

```
# Normally we would obtain these through a geocoding service
```

```
coordinates = {
```

```
    "Rennes": [48.1173, -1.6778],
```

```
    "Paris": [48.8566, 2.3522],
```

```
    "Saclay": [48.7290, 2.1912],
```

```
    "Lille": [50.6292, 3.0573],
```

```
    "Bordeaux": [44.8378, -0.5792],
```

```
    "Sophia": [43.6222, 7.0500], # This is an approximation for Sophia Antipolis
```

```
    "Nancy": [48.6921, 6.1844]
```

```
}
```

```
# Initialize a map centered around France
```

```
map_france = folium.Map(location=[46.603354, 1.888334], zoom_start=6)
```

```
# Add markers for each team
```

```
add_team_marker(map_france, "DIVERSE", "Rennes")
```

```
add_team_marker(map_france, "PICUBE", "Paris")
```

```
add_team_marker(map_france, "ARGO", "Paris")
```

```
add_team_marker(map_france, "DEDUCTEAM", "Saclay")
```

```
add_team_marker(map_france, "EVEREF", "Lille")
```

```
add_team_marker(map_france, "FLOWERS", "Bordeaux")
```

```
add_team_marker(map_france, "MNEMOSYNE", "Bordeaux")
```

```
add_team_marker(map_france, "SPIRALS", "Lille")
```

```
add_team_marker(map_france, "STAMP", "Sophia")
```

```
add_team_marker(map_france, "SED", "Nancy")
```

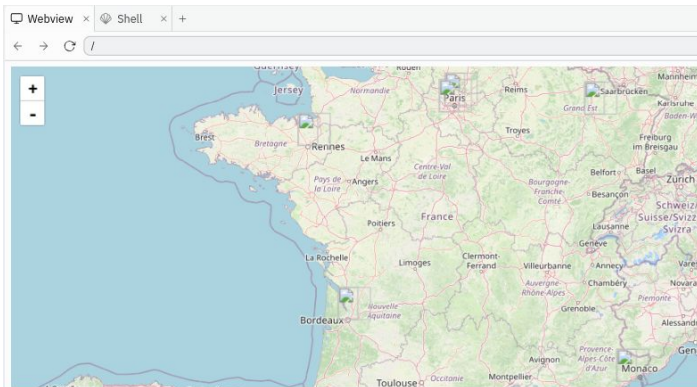
```
# Since we cannot display the folium map directly in this environment,  
# we'll save it to an HTML file which you can then open in a web browser.
```

```
map_file_path = '/mnt/data/france_teams_map.html'
```

```
map_france.save(map_file_path)
```

```
map_file_path
```

Please write a JavaScript program, using Leaflet.js



ChatGPT

Creating a JavaScript program using Leaflet.js to achieve a similar functionality as your Python code with Folium is a good way to visualize geographic data on web platforms. Leaflet is a leading open-source JavaScript library for mobile-friendly interactive maps.

Here's a basic example of how you could write such a program:

```
index.html x scripts.js x  
script.js > f addCityMarker > x popupText  
1 var map = L.map('map').setView([46.603354, 1.888334], 6);  
2  
3 // Add OpenStreetMap tiles:  
4 L.tileLayer('https://{s}.tile.openstreetmap.org/{z}/{x}/{y}.png', {  
5   maxZoom: 19,  
6   attribution: '© OpenStreetMap contributors'  
7 }).addTo(map);  
8  
9 // Teams and their respective cities  
10 var teams = [  
11   { city: "Rennes", team: "DIVERSE" },  
12   { city: "Paris", team: "PICUBE" },  
13   { city: "Paris", team: "ARGO" },  
14   { city: "Saclay", team: "DEDUCTEAM" },  
15   { city: "Lille", team: "EVEREF" },  
16   { city: "Bordeaux", team: "FLOWERS" },  
17   { city: "Bordeaux", team: "MNEMOSYNE" },  
18   { city: "Lille", team: "SPIRALS" },  
19  
20 index.html x scripts.js x  
21 }  
22 <!DOCTYPE html>  
23 <html>  
24 <head>  
25 <title>Team Locations in France</title>  
26 <link rel="stylesheet" href="https://unpkg.com/leaflet@1.0.0/dist/leaflet.css" />  
27 <script src="https://unpkg.com/leaflet@1.0.0/dist/leaflet.js"></script>  
28 <div id="map" style="height: 400px; width: 100%; border: 1px solid #ccc; margin: 0 auto; text-align: center; font-size: 24px; color: #000080; font-weight: bold; padding: 5px 0 0 0;">Map  
29 }  
30 </div>  
31 </body>  
32 </html>  
33  
34  
35  
36  
37  
38  
39  
40 // Helper function to get coordinates of a city  
41 function getCoordinates(city) {  
42   var coordinates = {  
43     "Rennes": [48.1173, -1.6778],  
44     "Paris": [48.8566, 2.3522],  
45     "Saclay": [48.7290, 2.1912],  
46     "Lille": [50.6292, 3.0573],  
47     "Bordeaux": [44.8378, -0.5792],  
48     "Sophia": [43.6222, 7.0500],  
49     "Nancy": [48.6921, 6.1844]  
50   };  
51   return coordinates[city];  
52 }  
53  
54  
55  
56  
57  
58
```

SE task: modernization/migration

```
add markers for each city  
Object.keys(teamsByCity).forEach(function(city) {  
  addCityMarker(city, teamsByCity[city]);  
});  
}
```

MA

You

I have the following Python code:

import folium

Given coordinates for the cities (these are example coordinates and may not be accurate)

Normally we would obtain these through a geocoding service

coordinates = {

"Rennes": [48.1173, -1.6778],

"Paris": [48.8566, 2.3522],

"Saclay": [48.7290, 2.1912],

"Lille": [50.6292, 3.0573],

"Bordeaux": [44.8378, -0.5792],

"Sophia": [43.6222, 7.0500], # This is an approximation for Sophia Antipolis

"Nancy": [48.6921, 6.1844]

}

Initialize a map centered around France

map_france = folium.Map(location=[46.603354, 1.888334], zoom_start=6)

Add markers for each team

add_team_marker(map_france, "DIVERSE", "Rennes")

add_team_marker(map_france, "PICUBE", "Paris")

add_team_marker(map_france, "ARGO", "Paris")

add_team_marker(map_france, "DEDUCTEAM", "Saclay")

add_team_marker(map_france, "EVEREF", "Lille")

add_team_marker(map_france, "FLOWERS", "Bordeaux")

add_team_marker(map_france, "MNEMOSYNE", "Bordeaux")

add_team_marker(map_france, "SPIRALS", "Lille")

add_team_marker(map_france, "STAMP", "Sophia")

add_team_marker(map_france, "SED", "Nancy")

Since we cannot display the folium map directly in this environment,

we'll save it to an HTML file which you can then open in a web browser.

map_file_path = '/mnt/data/france_teams_map.html'

map_france.save(map_file_path)

map_file_path

Please write a JavaScript program, using Leaflet.js

GP

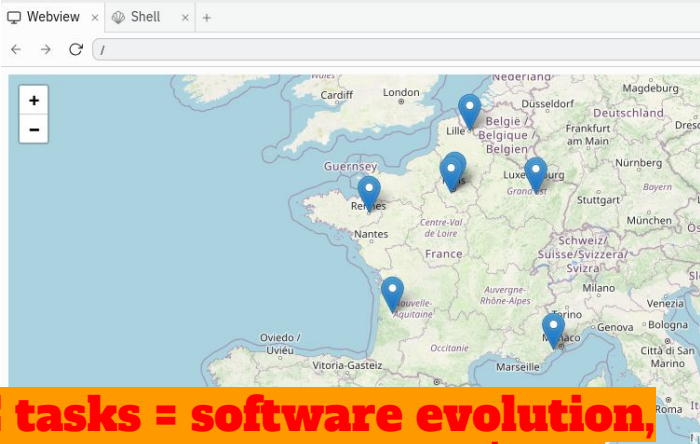
ChatGPT

Creating a JavaScript program using Leaflet.js to achieve a similar functionality as your Python code with Folium is a good way to visualize geographic data on web platforms. Leaflet is a leading open-source JavaScript library for mobile-friendly interactive maps.

Here's a basic example of how you could write such a program:

Polyglot?

- it's not a pure PL translation problem (it's also Folium => Leaflet)
- testing oracle problem
- PL and libraries evolve



SE tasks = software evolution, maintenance, upgrade | bug fixing | program understanding

```

index.html x  script.js x  +
script.js > f addCityMarker > X popupText
1  var map = L.map('map').setView([46.603354, 1.888334], 6);
2

index.html x  script.js x  +
index.html
1  <!DOCTYPE html>
2  <html>
3
4  <head>
5  <title>Team Locations in France</title>
6  <link rel="stylesheet" href="https://unpkg.com/leaflet@1.1.0/dist/leaflet.css" />
7  <script src="https://unpkg.com/leaflet@1.1.0/dist/leaflet.js"></script>
8  </head>
9  <style>
10  #map {
11    height: 400px;
12  }
13 </style>
14
15 <body>
16 <div id="map"></div>
17 <script src="script.js"></script>
18 </body>

29 10  height: 400px;
30 11  }
31 12  </style>
32 13  </head>
33 14  /
34 15  <body>
35 16  <div id="map"></div>
36 17  <script src="script.js"></script>
37 18  </body>
38 19
39 20  </html>
40
41 // Helper function to get coordinates of a city
42 function getCoordinates(city) {
43   var coordinates = {
44     "Rennes": [48.1173, -1.6778],
45     "Paris": [48.8566, 2.3522],
46     "Saclay": [48.7290, 2.1912],
47     "Lille": [50.6292, 3.0573],
48     "Bordeaux": [44.8378, -0.5792],
49     "Sophia": [43.6222, 7.0500],
50     "Nancy": [48.6921, 6.1844]
51   };
52   return coordinates[city];
53 }
54
55 // Add markers for each city
56 Object.keys(teamsByCity).forEach(function(city) {
57   addCityMarker(city, teamsByCity[city]);
58 });

```

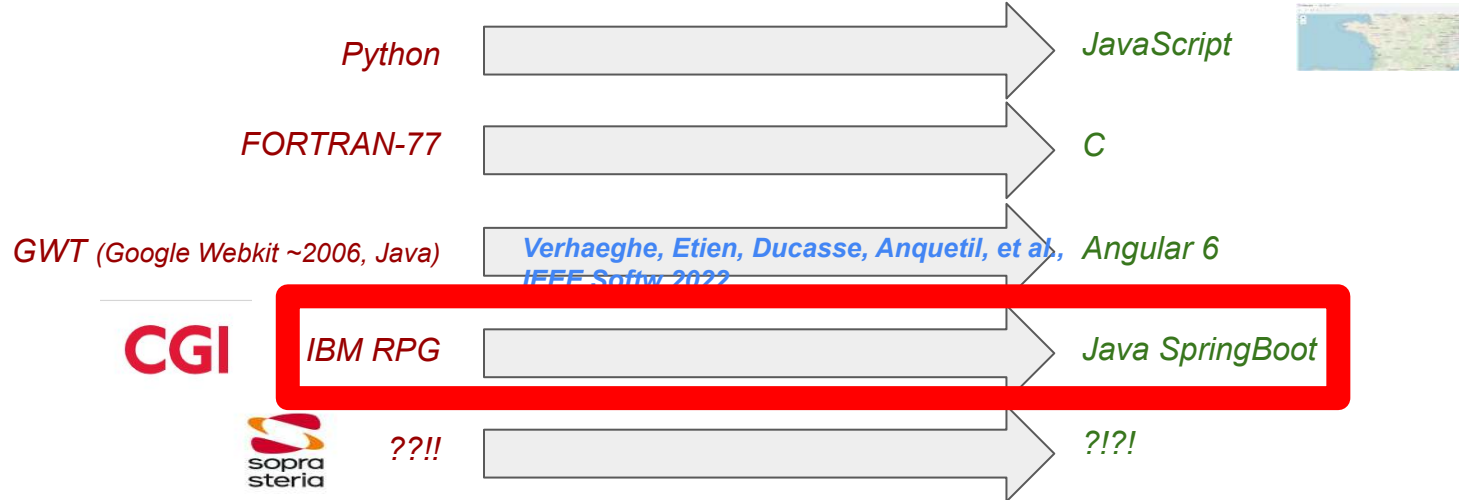
Software migration and modernization

Evref+DiverSE

Goal: LLMs to assist in migrating an existing/legacy code base to newer technologies

Approach:

- LLMs for reverse engineering, assistance to devs/architects
- Validation of the migration (testing), etc.
- Fine-grained examples and data out of compilers/transpilers
- Use-cases from Industrial collaborations



N'Guessan Hermann Kouadio (PhD student, CIFRE with CGI)



PhD advisors: Olivier Barais, Mathieu Acher, Jean-Marc Jézéquel (DiverSE)

Ongoing work: not directly in the défi, but highly related!

Opéra Desjardins x TSI AV SPD001A

RESSOURCES HUMAINES 19/10/07

IDENTIFICATION DE LA COMPAGNIE 15:32:52

No entreprise: Cle de référence :

No compagnie : LCOA1579 Date création : 2005/02/10

Desc. Longue : COMPAGNIE DE LISE LCOA1579 Abrégée : LISE LCOA1579

No / Rue : 111-475 AUBRY Code postal : J7R 7K2

Ville : ST-JEROME Province : QUE

Téléphone : 150 289-8384 Télécopieur : 500 005-9999

Courriel : vExhB1CwUsgZ00aBPH#fLjriAVwS.Dq

Nom légal : COMPAGNIE DE LISE LCOA1579

Secteur économique : PROMOT.EN CQ Rapport de base : 1 PAPIER

Statut compagnie :

Limite taux horaire : Site brute par employé : 5000.00

Intégrateur : XNLE Leger Nathalie Paie-maitre: INES Cardinal Lise

Mode saisie : 1 EMPD TELEP Date début : 2005/02/14 Niveau com.: A 30

Jour saisie : 2 Lundi Jour trait.: 2 Lundi

Hre appel : 10:12 Langue cle : F FRANCAIS Date transf: 2019/05/03

Civ. Contact Lg Tél. principal Poste Tél. secondaire

Mme TERIBU ENDAJH F 1 350 140-2141 1 236 163-2709

M REIIURO ERDAN F 1 392 963-0238 1 843 484-2881

F3=Sortir F4=Invite F9=Commentaire F10=Suivant F12=Précédent F24=Autres

Limite de défilement atteinte.

21/020

From

To

Employeur D TAVR OFFAC DOMINIC TEL EMPLOY PAIE OUTS Outils Outils

02/19/2025 09:00:00 (03/10/2025)

Conciliation bancaire

(03/10/2025)

Calendrier des chèques en circulation

Tableau des chèques en circulation

OPTION	# CHQ	# CIE	# MATR.	DT PAYABLE	MT CHQ	NOM MATR.
1	7326376	SP0A0954	00001245	2022/06/30	1.59	Larivée Vincent
2	7372062	SP0A0954	00000305	2022/06/30	104.50	RAMAIBOURIS MAKAYLA
3	7384	SP0A0954	00000270	2022/11/10	1.76	Yezina Thérèse
4	7394	SP0A0954	00000184	2022/11/10	964.84	St-Laurent Adèle
5	7395750	SP0A0954	00000302	2022/11/10	190.87	Galibert Julien
6	7395167	SP0A0958	00000834	2022/11/17	6.47	SOLAI AHAN YOU FAF YOU WAFID
7	7395168	SP0A0958	00000837	2022/11/17	3.64	ACHVIBI AICOM
9	7395153	SP0A0950	00000309	2022/11/17	160.26	KATI AKYRAN RAMMAYI
0	7395164	SP0A0958	00000830	2022/11/17	197.48	KA HANA VAN HANMEYAH
10	7406650	SP0A0950	00000289	2022/12/01	240.56	Djimon Mohammed Ibrahim
11	7406661	SP0A0958	00000842	2022/12/01	217.4	JACQUES SAMUEL
12	7406664	SP0A0950	00000308	2022/12/01	151.50	DEI FRIN FRANCOIS

Précédent 3 Suivant

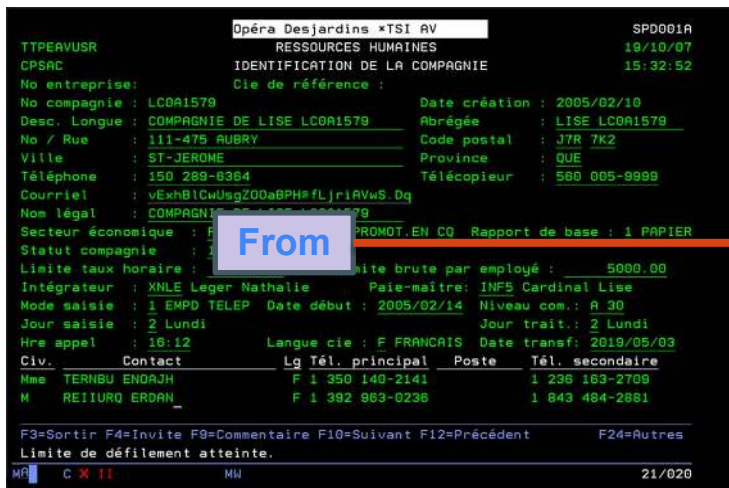
F3=Sortir F5=Rechercher F7=Tri (Asc) F12=Précédent F17=Rechercher F18=Tri (Des) F24=Autres

From a legacy application:

- developed with **RPG/AS400** language since **1990**
- executed on **Application System/400** (an IBM's OS)
- with a **terminal user interface**
- stored data on **DB2** database without explicit relation
- Up to **2 millions** lines of code (**LOC**)

To a modernized application:

- **3-tiers** architecture
- running on **Linux**
- backend in **Java (Spring boot)** and front-end in **Angular**
- **web interface**
- a relational database: **SqlServer**



The screenshot shows a modern web application interface for 'Conciliation bancaire'. It features a table with columns: OPTION, # CHQ, # CIE, # MATR., DT PAYABLE, MT CHQ, and NOM MATR. An orange arrow points from the 'From' box in the legacy application to a blue box with the word 'To' overlaid on the table. The table lists 12 rows of transaction data.

OPTION	# CHQ	# CIE	# MATR.	DT PAYABLE	MT CHQ	NOM MATR.
1	7328335	SPMA094	00001245	2022/06/30	1,59	Larivée Vincent
2	7372092	SPMA095	00001005	2022/06/30	104,50	RAMBOURS MAKAYLA
3	7381	SPMA096	00001270	2022/11/10	1,76	Veina Thérèse
4	7591	SPMA097	00001884	2022/11/10	864,54	Silouan Adèle
5	7395750	SPMA098	00003012	2022/11/10	190,57	Galibert Julien
6	7395167	SPMA099	00003834	2022/11/17	6,47	SOUAI AHANOU IAH-SOU BIYED
7	7391168	SPMA100	00003697	2022/11/17	3,64	ACHVIE ALIOM
8	7391153	SPMA101	00003019	2022/11/17	169,26	KATI AKRAN RAMMCYLI
9	7391164	SPMA102	00003829	2022/11/17	19,748	KA HANA VAN HANMEYAH
10	7406650	SPMA103	00003989	2022/12/01	240,56	Djavan Mohammed Ibrahim
11	7406661	SPMA104	00003942	2022/12/01	21,14	JACQUES SAMUEL
12	7406674	SPMA105	00003938	2022/12/01	151,50	DEFRIN FRANCOISE

The current migration is done using deterministic tools (lexer/parser + IR)
Currently, **over 70%** of the code is migrated with this process

The current migration is done using deterministic tools (lexer/parser + IR)

Currently, **over 70%** of the code is migrated with this process

Difficulties:

- long methods (up to 10K lines!)
- methods are not modularized (calls to decompose and refactor)
- RPG variants
- context matters:
 - hard to only translate one method (you need to take the context into account)
 - implicit semantics (eg with variable' names)
 - different RPG files (physic: table; logic: view)
- reading and code comprehension

```
C      Read      CLE001
C      Eval      MsgId = 'CLR*ALL'
C      ExSr      SndMsg
C
C      Eval      *In12 = *Off
C      Eval      *In51 = *On
C
C      Select
C
C      When      *InKc Or *InKl
C      KC        MOVE      'C'          #CDRET
C      KL        MOVE      'L'          #CDRET
C      Leave
C      When      *InKj
C      ExSr      InzSfl
C      If      SwitchF10 = *On
C      C      Eval      SwitchF10 = *Off
C      C      ExSr      SrSfl2
C
C      Else
C      C      Eval      SwitchF10 = *On
C      C      ExSr      SrSfl
C      EndIf
C
C      When      *In56
C
C      4      Chain      Sfl001
C      If      #DtApd <>0
C
C      CALL      'RHU815R'
C      PARM      '*VDATE'          ACT          8
C      PARM      *ZEROS          DEBDAT       8 0
C      PARM      *ZEROS          FINDAT       8 0
C      PARM      *ZEROS          CALCUL       6 0
C      PARM      #dtapd          DATREF       8 0
C      PARM      *ZEROS          DEFS         8 0
```

RPG source code. Src: Opera project sreenshot

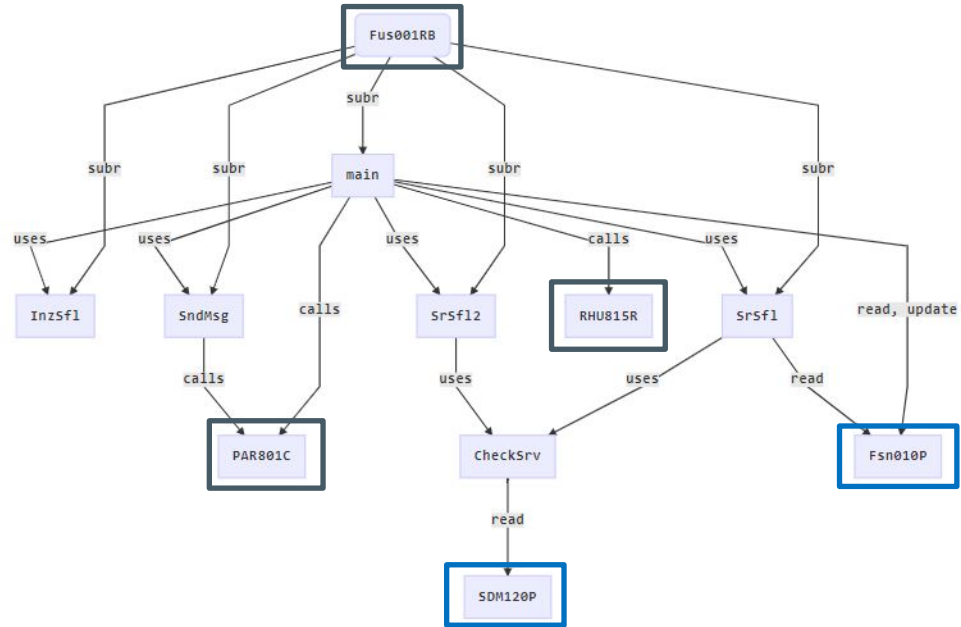
Without documentation!

REVERSE ENGINEERING

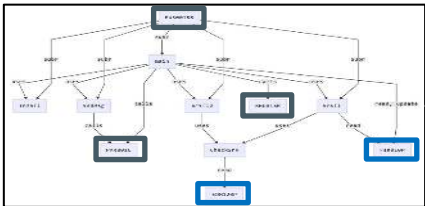


Augmenting the context before calling LLMs

Build the program tree execution flow for retro-documentation



RPG source code



Prompt Temperature ...

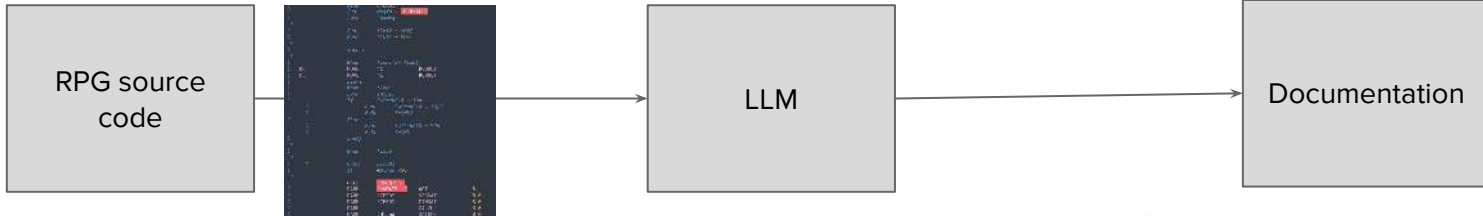
LLM

Documentation

This RPG method, named 'CheckSrv', is designed to verify the presence of a report named 'SPD616' in a database entity called 'Spdm1211'. The method starts by setting a flag 'FoundSpd616' to '*Off'. Then, it checks if the report exists in the database by evaluating the contents of 'Spdm1211'. If the report is found and not at the end of the database, the flag 'FoundSpd616' is set to '*On'.

Can you generate the most explicit summary possible of this method with as few words as possible for a developer ?
Try to replace technical names by the descriptions of elements given in your global context.
If a file is used, use this description to explain the code, and refer to the technical name in parentheses.
Don't say "in the logical file xxx" with xxx is the technical name, use the description instead.
Don't refer to unused context. Don't talk about what the code doesn't contain.
Provide two summaries in this order:
- general one into #general# tag: must be short
- detailed one into #details# tag: must be more explicit

You are a developer expert in mainframe language and in RPG/AS400 with IBM-i and you will analyze code content for us.
You will summarize code block in RPG/AS400 running IBM-i mainframe platform.



Evaluation of LLM pipeline wrt retro-documentation

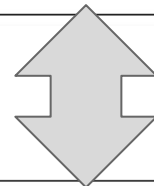
Possible approaches:

- manually review generated doc
- compare GT doc with generated doc
 - Levenshtein distance? BLEU? ROUGE?

We aim to know:

- precision
- recall
- in an automated way

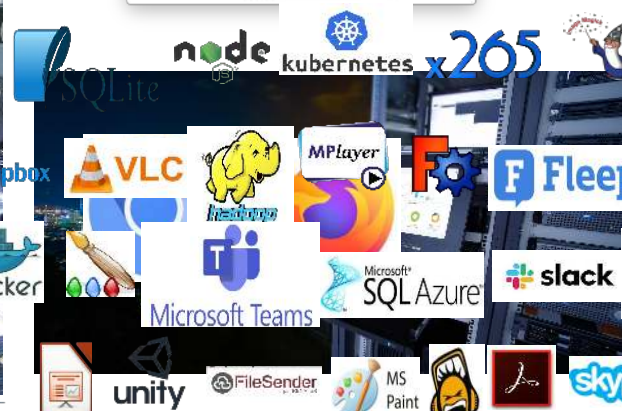
This RPG method, named 'CheckSrv', is designed to verify the presence of a report named 'SPD616' in a database entity called 'Spdm121L'. The method starts by setting a flag 'FoundSpd616' to '*Off'. Then, it checks if the report exists in the database by evaluating the contents of 'Spdm121L'. If the report is found and not at the end of the database, the flag 'FoundSpd616' is set to '*On'.



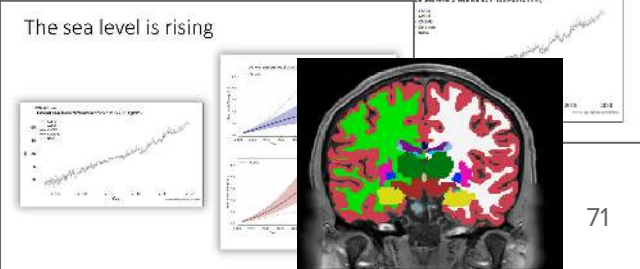
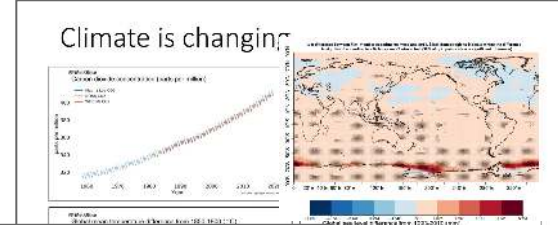
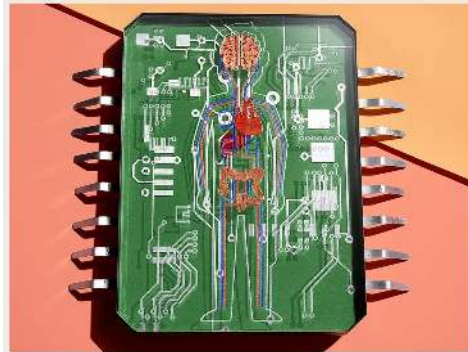
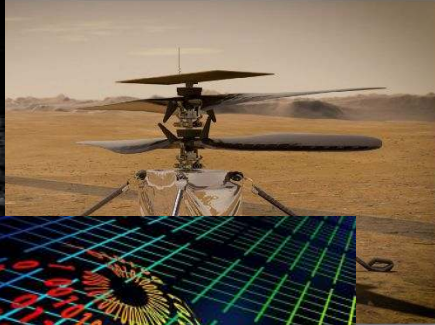
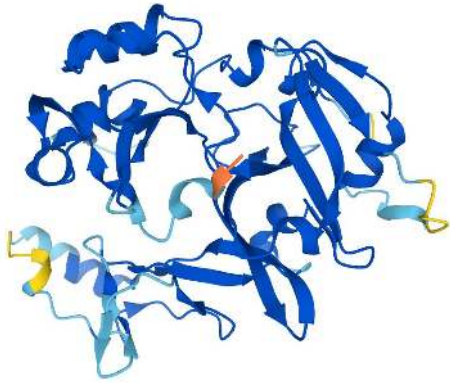
Ground truth documentation

- Reading the "Service agreement, items to be invoiced" data
 - Checks that the data is already present
 Verifies that the service agreement table (Spdm121L) contains an item associated with the company number and the service code (code = SPD616). If the searched item is present, FoundSpd616 is set to true.

Software is eating the world!



Software is eating science!

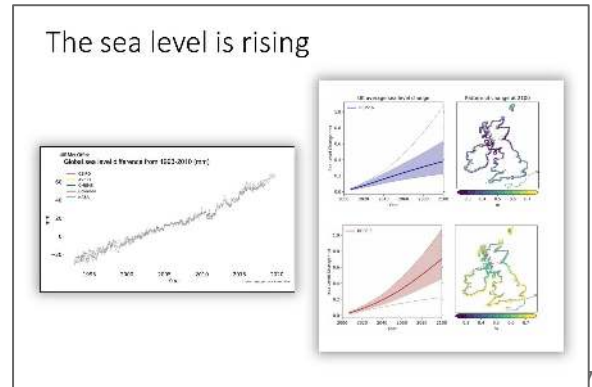
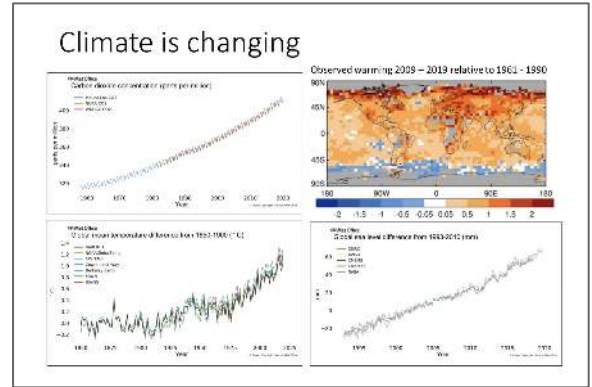


Computational science depends on software and its engineering



design of mathematical model
mining and analysis of data
executions of large simulations
problem solving
executable paper

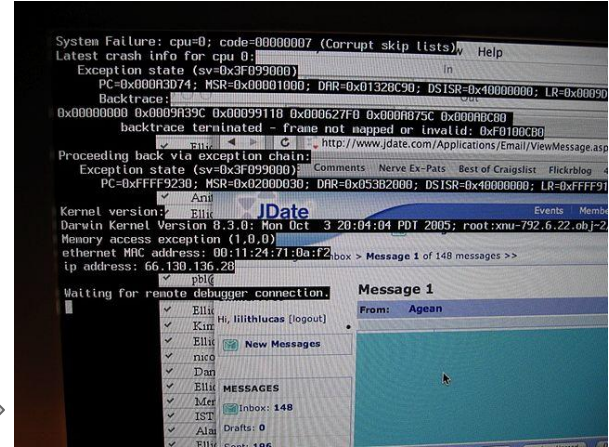
from a set of scripts to automate the deployment to... a comprehensive system containing several features that help researchers exploring various hypotheses



Computational science depends on software and its engineering



multi-million line of code base
multi-dependencies
multi-systems
multi-layer
multi-version
multi-person
multi-variant



```
In [6]: sess.run(tf.svd(tf_matrix))  
Out[6]: (array([[ 9.9998987e-01,  1.48747023e-03,  4.88133628e-06,  
  4.69811084e-06,  4.37980998e-06,  3.45290823e-06,  
  1.14686304e-06,  3.10980795e-06,  2.97625912e-06,  
  2.65099743e-06,  1.91537106e-06,  0.00000000e+00,  
  0.00000000e+00,  0.00000000e+00,  0.00000000e+00,  
  0.00000000e+00,  0.00000000e+00,  0.00000000e+00,  
  0.00000000e+00,  0.00000000e+00,  0.00000000e+00,  
  0.00000000e+00,  9.82503479e-05, -2.52892733e-06,  
  6.43756945e-07, -2.61201406e-06, -3.18651830e-06,  
  0.00000000e+00,  nan, nan, nan,  
  8.60062854e-10,  1.93595340e-09,  0.00000000e+00,  
 -1.24836730e-09,  3.83645737e-09, -3.90316446e-09,  
  nan, -7.00323994e-07], dtype=float32),  
 [ -7.27595761e-11,  7.15255737e-07, -2.68207733e-02,  
 -6.68754578e-01,  1.68675050e-01, -2.37232931e-02, ...])
```

Dealing with software collapse: software stops working eventually
Konrad Hinsien 2019
Configuration failures represent one of the most common types of
software failures Sayagh et al. TSE 2018

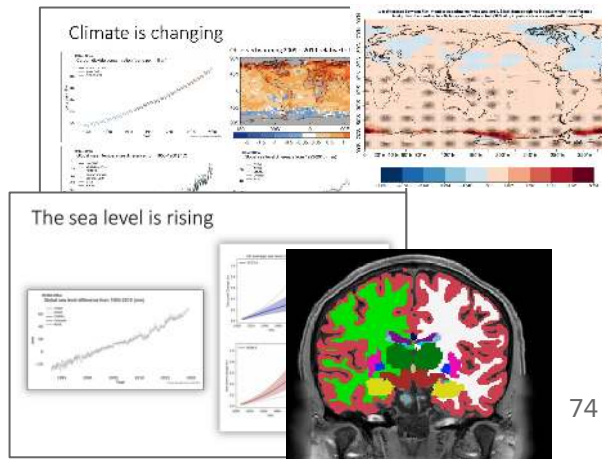
Reproducible science

“Authors provide all the necessary data and the computer codes to run the analysis again, re-creating the results.”

Despite the availability of data and code, several studies report that the same data analyzed with different software can lead to different results.



from a set of scripts to automate the deployment to... a comprehensive system containing several features that help researchers exploring various hypotheses



As **Software is key to Computational Science and Society**, numerous challenges and opportunities ahead!

- democratization of software
- reliable/trustable/resilient software
- speeding up productivity (and focusing on tough/creative challenges)
- software to explore variants' space (hypothesis, ideas, products, etc.)
- ...

We need to **invest more in software science and software engineering (SE) research**

Large Language Models (LLMs) are a transformative technology to raise the level of abstraction, automate SE tasks, and augment/assist software users/developers/scientists → **LLM for SE!?**

Generative AI and LLM for Software



starcoder2

Public



GitHub Copilot



```

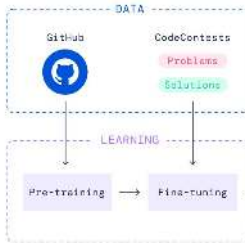
// [[X]] [[Y]] [[Z]] [[W]]
let h = (x, y, z, w) => {
  return x + y + z + w;
};

// [[X]] [[Y]] [[Z]] [[W]]
let h = (x, y, z, w) => {
  return x + y + z + w;
};
    
```

```

// [[X]] [[Y]] [[Z]] [[W]]
let h = (x, y, z, w) => {
  return x + y + z + w;
};

// [[X]] [[Y]] [[Z]] [[W]]
let h = (x, y, z, w) => {
  return x + y + z + w;
};
    
```



1553_D_Backspace
python
pass "Layer 18" [Up]

123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

Problem Description

1553_D_Backspace

python

pass "Layer 18" [Up]

Solution

```

// [[X]] [[Y]] [[Z]] [[W]]
let h = (x, y, z, w) => {
  return x + y + z + w;
};

// [[X]] [[Y]] [[Z]] [[W]]
let h = (x, y, z, w) => {
  return x + y + z + w;
};
        
```

Llama-2



DeepSynth



AlphaCode

Code synthesis: default strategy to answer questions...



You

what is the result of 2+3 ?



ChatGPT



Finished analyzing ▾

```
python
```

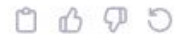
Copy code

```
# Calculating the sum of 2 and 3  
result = 2 + 3  
result
```

Result

5

The result of $2 + 3$ is 5. [↩]



Related work (code generation multi language)

MultiPL-E: A Scalable and Polyglot Approach to Benchmarking Neural Code Generation

Federico Cassano et al. TSE 2023

Piloting Copilot and Codex: Hot Temperature, Cold Prompts, or Black Magic?

Jean-Baptiste Döderlein, Mathieu Acher, Djamel Eddine Khelladi, Benoit Combemale, 2022, <https://arxiv.org/abs/2210.14699>

Evaluation on HumanEval or
**Mainly at the methods level, from a specification
Discrepancies across PL**

Model	Size	Multi-lingual Human-Eval					
		C++	Java	PHP	TS	C#	Bash
CodeGen-Multi	16B	21.0%	22.2%	8.4%	20.1%	8.2%	0.6%
CodeGeeX	13B	16.9%	19.1%	13.5%	10.1%	8.5%	2.8%
code-cushman-001	12B	30.6%	31.9%	28.9%	31.3%	22.1%	11.7%
StarCoder Base	15.5B	30.6%	28.5%	26.8%	32.2%	20.6%	11.0%
StarCoder Python	15.5B	31.6%	30.2%	26.1%	32.3%	21.0%	10.5%
LLAMA-v2	7B	6.8%	10.8%	9.9%	12.6%	6.3%	3.2%
	13B	13.7%	15.8%	13.1%	13.2%	9.5%	3.2%
	34B	23.6%	22.2%	19.9%	21.4%	17.1%	3.8%
	70B	30.4%	31.7%	34.2%	15.1%	25.9%	8.9%
CODE LLAMA	7B	28.6%	34.2%	24.2%	33.3%	25.3%	12.0%
	13B	39.1%	38.0%	34.2%	29.6%	27.3%	15.2%
	34B	47.8%	45.6%	44.1%	33.3%	30.4%	17.1%
	70B	52.8%	51.9%	50.9%	49.1%	38.0%	29.1%

	Copilot pass@1	Codex pass@1
Original (HumanEval, Python3)	31.1	22.44
Original (Leetcode, Javascript)	37.3	32.3
Original (Leetcode, Java)	46.3	39.3
Original (Leetcode, C#)	44.3	33.7
Original (Leetcode, Python3)	33.7	30.3
Original (Leetcode, C)	18.7	18.3
Original (Leetcode, C++)	45.3	41.3

TABLE II: Pass@1 (original prompts; default temperature) 78

Related work (code generation multi language)

Lost in Translation : A Study of Bugs Introduced by Large Language Models while Translating Code
Rangeet, Pan, et al. ICSE 2024

ICSE 2024, April 2024, Lisbon, Portugal

Pan et al.

Table 2: Performance of subject LLMs in translating code from different studied datasets. The best performance by general and code LLMs are highlighted in teal and violet, respectively. The final performance is computed over the average of each dataset.

Dataset	Source Language	Source Samples	#Tests	Target Language	#Translations	% Successful Translations						
						CodeGen	CodeGeeX	StarCoder	GPT-4	Llama 2	TB-Airoboros	TB-Vicuna
CodeNet [63]	C	200	200	C++, Go, Java, Python	800	23.4%	14.9%	42.0%	83.0%	14.9%	18.8%	4.4%
	C++	200	200	C, Go, Java, Python	800	14.0%	3.6%	39.1%	80.0%	9.5%	8.3%	3.4%
	Go	200	200	C, C++, Java, Python	800	14.3%	5.9%	42.0%	85.5%	16.9%	6.6%	0.9%
	Java	200	200	C, C++, Go, Python	800	21.3%	10.3%	30.3%	81.3%	13.9%	6.5%	0.1%
	Python	200	200	C, C++, Go, Java	800	17.5%	7.3%	33.3%	79.9%	11.0%	6.5%	1.0%
Total/Average (CodeNet)	-	1,000	1,000	-	4,000	18.1%	8.4%	37.3%	82.0%	13.2%	9.3%	2.0%
AVATAR [23]	Java	249	6,255	C, C++, Go, Python	996	8.1%	1.8%	11.9%	70.8%	1.8%	5.1%	0.0%
	Python	250		C, C++, Go, Java	1,000	3.8%	1.6%	14.2%	52.2%	4.7%	0.9%	0.9%
Total/Average (AVATAR)	-	499	6,255	-	1,996	5.9%	1.7%	13.0%	61.5%	3.2%	3.0%	0.4%
EvalPlus [54]	Python	164	2,682	Java	164	16.5%	3.7%	22.0%	79.3%	1.2%	14.0%	7.9%
Commons CLI [6]	Java	22	310	Python	22	0.0%	0.0%	0.0%	13.6%	0.0%	0.0%	0.0%
Click [10]	Python	15	611	Java	15	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Total/Average (All)	-	1,700	10,858	-	6,197	8.1%	2.8%	14.5%	47.3%	3.5%	5.3%	2.1%

Many difficulties to translate real projects from one programming language to another one

Challenges and Opportunities for Science and Society

Challenges:

- LLMs are:
 - brittle, leading to faulty, under-optimized, or unreliable code (including tests!)
 - under-trained on “small” languages or not up-to-date with (knowledge of) new libraries or software versions
 - LLMs lack planning and reasoning capabilities
- Challenges of improving LLMs:
 - Better “language” than prompt?
 - more verification/validation tools to verify the outcome?
 - fine-tuning?
- Benchmarking (SE task specific, no data contamination, etc.)
- Open source LLMs (see next slides!)



LLM4Code: Reliable and Productive Code Assistants
Based on Large Language Models

Project Leads

DIVERSE	Reims	Mathieu Acher	mathieu.acher@inria.fr
PICUBE	Paris	Guillaume Baudart	guillaume.baudart@inria.fr

Internal Partners (EP/CRI)

ARGO	Paris	Marc Lelarge	marc.lelarge@inria.fr
DEDUCTEAM	Saclay	Anthony Bordg	anthony.bordg@inria.fr
EVREF	Lille	Anne Etien	anne.etien@univ-lille.fr
FLOWERS	Bordeaux	Pierre-Yves Oudeyer	pierre-yves.oudeyer@inria.fr
MNEMOSYNE	Bordeaux	Xavier Hinaut	xavier.hinaut@inria.fr
SPIRALS	Lille	Clément Quinton	clement.quinton@inria.fr
STAMP	Sophia	Cyril Cohen	cyril.cohen@inria.fr
SED	Nancy	Laurent Pierron	laurent.pierron@inria.fr

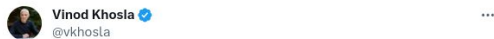
External Partners

SOFTWARE HERITAGE	Roberto di Cosmo	roberto@dicosmo.org
LABRI	Nathanaël Fijalkow	nathanael.fijalkow@labri.fr
	Romain Robbes	romain.robbes@cnrs.fr
SOPRA STERIA	Yann Gloriant	yann.gloriant@soprasteria.com

Opportunities:

- democratization of software
- reliable/trustable/resilient software
- speeding up productivity (and focuses on hard/creative tasks)
- software to explore variants' space (hypothesis, ideas, products, etc.)
- ...

Software or/and ML are eating the world...



Portfolio companies @OpenAI and @replit set this off - a total paradigm shift. No surprise a marketing person just won a coding competition.

Prediction #2

There could be a billion programmers all "programming" in natural language. Many parts of the craft of coding would be commoditized & the entire field would be opened up to those without a CS degree. Democratizing computer coding is well on its way in the next 10-15 yrs.



At a hackathon where a winner is a "nontechnical" PM and her work — powered by Replit + AI — is more technically impressive than teams of engineers! Was surprised at first but it struck me that PMs must be exceptional prompting, after that "is" their job.



People often ask if ML or software skills are more the bottleneck to AI progress. It's the wrong question—both are invaluable, and people with both sets of skills can have outsized impact. We find it easier, however, to teach people ML skills as needed than software engineering.



coming up with good ml research ideas is significantly easier than implementing them in a complex codebase — the returns to being an extremely good engineer are super high



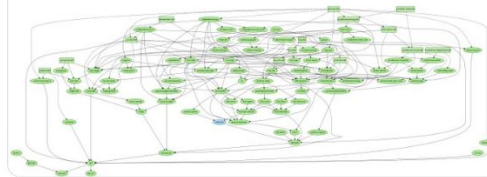
Update: it's finished now.



This is just wild!

Nov 9: @wtgowers, Green, Manners, and Tao prove the polynomial Freiman-Ruzsa conjecture in char 2. arxiv.org/abs/2311.05762

[Show more](#)



This is one of the hidden secrets of scalable research labs: creating template experimental codebases that allow fast testing of wild research ideas. You need your experiment, data, compute, and model abstractions just right 1/



Replying to @tszzl

It also follows that the returns to making simple codebases to test research ideas must be really high and also that the returns to making ai engineering tools are enormously high

Newcomers

End-users

Professional developers
(~28 millions)

Data analysts

Scientists
(physicists,
mathematicians,
historians, etc.)

As code is repetitive and predictable, there is hope to have tools that enhance reliability and productivity

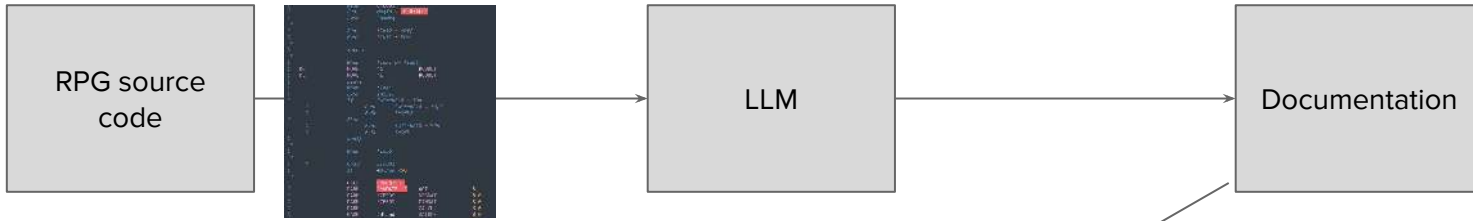
On the Naturalness of Software

Abram Hindle, Earl Barr, Mark Gabel, Zhendong Su, Prem Devanbu
devanbu@cs.ucdavis.edu

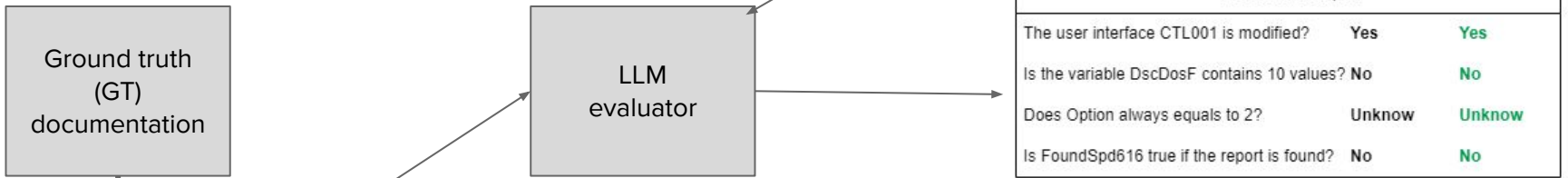
Unpublished version of ICSE 2012 paper, with expanded future work section
Enjoy! Comments Welcome.

Programming languages, in theory, are complex, flexible and powerful, but the programs that real people actually write are mostly simple and rather repetitive, and thus they have usefully predictable statistical properties that can be captured in statistical language models and leveraged for software engineering tasks.

Corpus-based n-gram model outperforms Eclipse **completion** for Java



Evaluation of LLM pipeline wrt retro-documentation

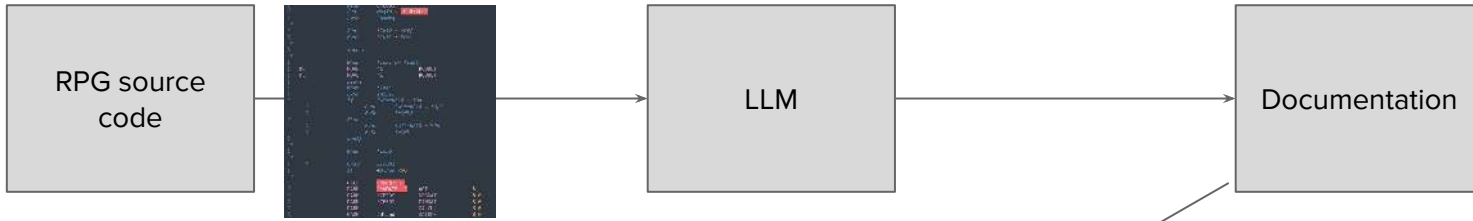


The user interface CTL001 is modified?	Yes	Yes
Is the variable DscDosF contains 10 values?	No	No
Does Option always equals to 2?	Unknown	Unknown
Is FoundSpd616 true if the report is found?	No	No

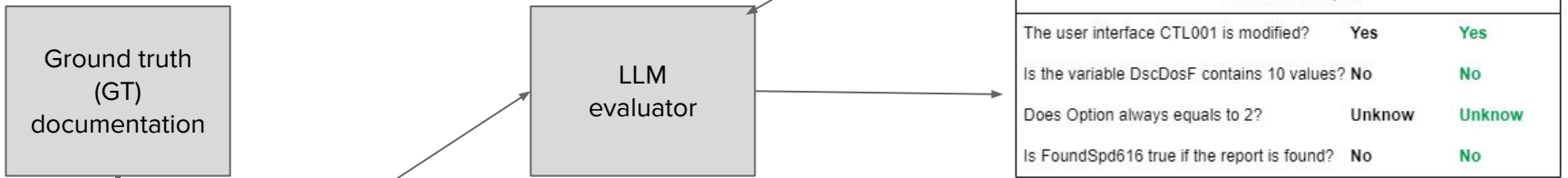
LLM evaluator answers to questions out of the generated documentation (instead of out of GT documentation)

Then we can compute precision/recall/accuracy/...

Questions with ground truth answers associated with GT documentation



Evaluation of LLM pipeline wrt retro-documentation



LLM evaluator answers to questions out of the generated documentation

Challenges:

- calibration/truth worthiness of the LLM evaluator
- documentation vs or with program comprehension
- granularity details of the doc

Questions with ground truth answers associated with GT documentation

Evaluation dataset

Our empirical evaluation is based on self-constructed dataset using previous rules to interpret

	program	subr	category	question	exp	prov
1	FUS001RB	InzSfl	General	The user interface CTL001 is modified ?	Yes	<null>
2	FUS001RB	InzSfl	General	Is the variable Desc_Dossier been modified base on the user language ?	Yes	<null>
3	FUS001RB	InzSfl	General	Is the variable Dsc_DosF contains many values ?	Yes	<null>
4	FUS001RB	InzSfl	General	Is the variable Dsc_DosF contains 10 values ?	No	<null>
5	FUS001RB	InzSfl	General	Is the current date set ?	No	<null>
6	FUS001RB	InzSfl	General	Is indicator IN60 on (active) before leave this subroutine ?	Yes	<null>
7	FUS001RB	InzSfl	General	Is indicator IN64 on (active) before leave this subroutine ?	No	<null>
8	FUS001RB	InzSfl	General	Do we wait the form submission to process data in the main subroutine ?	No	<null>
9	FUS001RB	InzSfl	OutOfBox	Do we read Dsc_DosF if the user language is English ?	No	<null>
10	FUS001RB	InzSfl	OutOfBox	Do we read Dsc_DosF if the user language is French ?	Yes	<null>
11	FUS001RB	*InzSr	General	Are the program input parameters defined?	Yes	<null>
12	FUS001RB	*InzSr	General	Is the current date is set ?	Yes	<null>
13	FUS001RB	*InzSr	General	Is indicator *in49 set following on variable LdaLge ?	Yes	<null>
14	FUS001RB	*InzSr	General	Does variable LdaLge equals 'A' means indicator *in49 false ?	No	<null>
15	FUS001RB	*InzSr	OutOfBox	Does Option always equals to 2 ?	Unknown	<null>
16	FUS001RB	*InzSr	OutOfBox	Is the key Key_Spdm121l used in main subroutine ?	Unknown	<null>
17	FUS001RB	*InzSr	OutOfBox	Does indicator in50 to on clear all data in Spdm121L file ?	Unknown	<null>
18	FUS001RB	CheckSrv	General	Is RH@044 a report ?	Unknown	<null>
19	FUS001RB	CheckSrv	General	Is FoundSpd616 true if the report is found?	Yes	<null>

More than **50 questions** for the **FUS001RB** program (average of 500 LOC)

Table 11. Distribution of SE tasks over six SE activities.

SE Activity	SE Task	Total	
Requirements engineering	Anaphoric ambiguity treatment (3)	Requirements term identification (1)	11
	Requirements classification (3)	Coreference detection (1)	
	Requirement analysis and evaluation (2)	Traceability automation (1)	
Software design	GUI retrieval (1)	Software specification synthesis (1)	3
Software development	Rapid prototyping (1)		136
	Code generation (62)	Agile story point estimation (1)	
	Code completion (16)	API documentation smell detection (1)	
	Code summarization (10)	API entity and relation extraction (1)	
	Code understanding (7)	Code optimization (1)	
	Code search (5)	Code example recommendation (1)	
	Program synthesis (5)	Control flow graph generation (1)	
	API recommendation (2)	Data analysis (1)	
	API synthesis (2)	Identifier normalization (1)	
	Code comment generation (2)	Instruction generation (1)	
	Code representation (2)	Type inference (1)	
	Method name generation (2)	Others (11)	
Software quality assurance	Test generation (8)	Bug localization (1)	24
	Vulnerability detection (7)	Failure-inducing test identification (1)	
	Test automation (4)	Flaky test prediction (1)	
	Verification (2)		
Software maintenance	Program repair (23)	Duplicate bug report detection (1)	58
	Code review (6)	Decompilation (1)	
	Debugging (4)	Program merge conflicts repair (1)	
	Bug report analysis (3)	Sentiment analysis (1)	
	Code clone detection (3)	Tag recommendation (1)	
	Logging (2)	Vulnerability repair (1)	
	Bug prediction (1)	Commit classification (1)	
	Bug triage (1)	Traceability recovery (1)	
	Bug report replay (1)	Others (6)	
Software management	Effort estimation (1)		1

• [Papers](#)

- [Type Inference](#)
- [Code Completion](#)
- [Code Generation](#)
- [Code Summarization](#)
- [Code Embeddings/Representation](#)
- [Code Changes/Editing](#)
- [Code Comments](#)
- [Bug/Vulnerability Detection](#)
- [Source Code Modeling](#)
- [Program Repair](#)
- [Program Translation](#)
- [Program Analysis](#)
- [Software Testing](#)
- [Code Clone Detection](#)
- [Code Language Models](#)
- [Code Review](#)
- [Code Documentation](#)
- [Empirical Studies](#)
- [Surveys](#)
- [Misc](#)

SE tasks!

Empirical experimentation results

SUBR	QUESTION	EXP	0						0,2						0,4					
			No-P		Lo-P		Fu-P		No-P		Lo-P		Fu-P		No-P		Lo-P		Fu-P	
			PROV	SCORE	PROV	SCORE	PROV	SCORE	PROV	SCORE	PROV	SCORE	PROV	SCORE	PROV	SCORE	PROV	SCORE	PROV	SCORE
SRSFL	Is the data recovery (CHAIN) in FND010P done on #Compt ?	No	No		No		No		No		No		No		No		No		No	
	Is the CheckSrv subroutine executed on the sixth iteration?	No	No		No		No		No		No		No		No		No		No	
	Is the data written to SFL001 coming from the FND010P file?	Yes	No		No		No		No		No		No		No		No		No	
	Is the data written to SFL001 coming from the Spdm1211 file?	No	No	57,14	No	57,14	No	71,43	No	57,14	No	57,14	No	57,14	No	57,14	No	71,43	No	71,43
	If the CHAIN on FND010P file returns a value, is the flag 90 set to true (=1)?	Yes	No		No		Yes		No		No		No		No		Yes		Yes	
SRSFL2	Is the SFL001 record written 7 times if FoundSpd616 is set on?	Yes	No		No		No		No		No		No		No		No		No	
	Is the data written to SFL001 coming from the Spdm1211 file?	No	No		No		No		No		No		No		No		No		No	
	Is the data recovery (CHAIN) in FND010P done on #Compt ?	No	No		No		No		No		No		No		No		No		No	
	Is the CheckSrv subroutine executed on the sixth iteration?	No	No		No		No		No		No		No		No		No		No	
	Is the data written to SFL001 coming from the FND010P file?	Yes	No	60	No	60	No	60	No	60	No	60	No	60	No	60	No	60	No	60
SndMsg	Is the SFL001 record written 7 times if FoundSpd616 is set on?	Yes	No		No		No		No		No		No		No		No		No	
	Is the data written to SFL001 coming from the Spdm1211 file?	No	No		No		No		No		No		No		No		No		No	
	Is PAR801C a program ?	Yes	No	50	Yes	100	Yes	100	No	50	Yes	100	Yes	100	No	50	Yes	100	No	0
CheckSrv	Is the PAR801C program called in this subroutine?	Yes	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
	Is RH@044 a report ?	No	No		No		No		No		No		No		No		No		No	
	Is FoundSpd616 true if the report is found?	Yes	Yes	75	Yes	50	Yes	50	Yes	75	Yes	75	Yes	75	Yes	50	Yes	50	Yes	75
	Is Spdm1211 data file read in this subroutine?	Yes	Yes		No		No		Yes		Yes		Yes		No		No		Yes	
	Does fields Cd@Srv always equals to 'SPD616'?	Yes	No		No		No		No		No		No		No		No		No	
InzSfl	The user interface CTL001 is modified ?	Yes	No		No		Yes		Yes		No		No		No		No		No	
	Is the variable Desc_Dossier been modified base on the user language ?	Yes	No		No		Yes		No		No		No		No		Yes		Yes	
	Is the variable Desc_Dossier contains many values ?	Yes	No		No		Yes		No		No		No		No		No		Yes	
	Is the variable Desc_Dossier contains 10 values ?	No	No	50	No	50	No	87,5	No	62,5	No	50	No	50	No	50	No	62,5	No	75
	Is the current date is set ?	No	No		No		No		No		No		No		No		No		No	
*InzSr	Do we wait the form submission to process data in the main subroutine ?	No	No		No		No		No		No		No		No		No		No	
	Do we read Dsc_DosF if the user language is English ?	No	No		No		No		No		No		No		No		No		No	
	Do we read Dsc_DosF if the user language is French ?	Yes	No		No		Yes		No		No		No		No		No		No	
	Are the program input parameters defined?	Yes	Yes		Yes		No		Yes		Yes		Yes		No		No		No	
	Is the current date is set ?	Yes	Yes	100	Yes	100	Yes	50	Yes	100	Yes	100	Yes	75	Yes	50	Yes	50	Yes	75
FUS001RB				63,33		63,33		70,00		66,67		66,67		63,33		53,33		63,33		66,67

Software migration and modernization

Evref+DiverSE

Goal: LLMs to assist in migrating an existing/legacy code base to newer technologies

Approach:

- LLMs for reverse engineering, assistance to devs/architects
- Validation of the migration (testing), etc.
- Fine-grained examples and data out of compilers/transpilers
- Use-cases from Industrial collaborations

Ongoing:

- recruitment of a PhD student
- SopraSteria in the défi

