



McGill

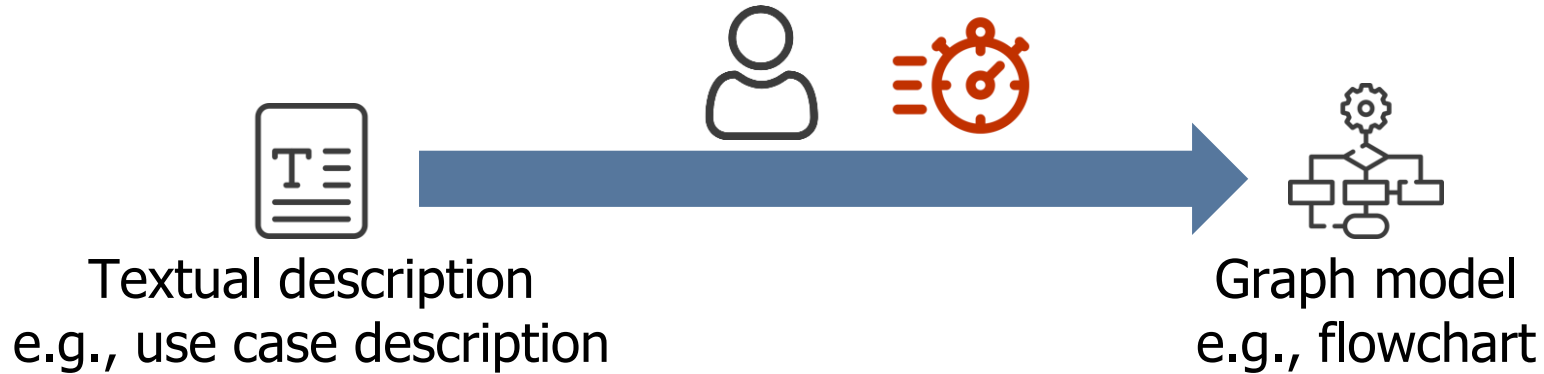


HUAWEI

Consistent Graph Model Generation with Large Language Models

Boqi Chen¹

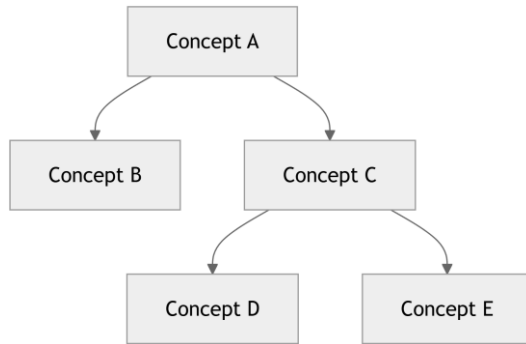
Manual Model Generation



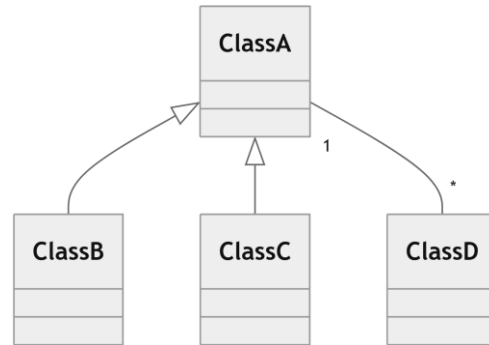
Manual generation can be **time-consuming** and **error-prone**

Graph Models

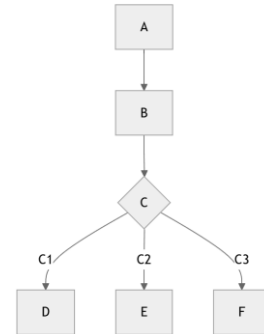
Many graph models are used in many SE processes



Taxonomies



Domain models

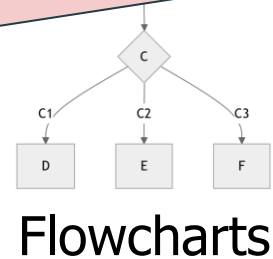
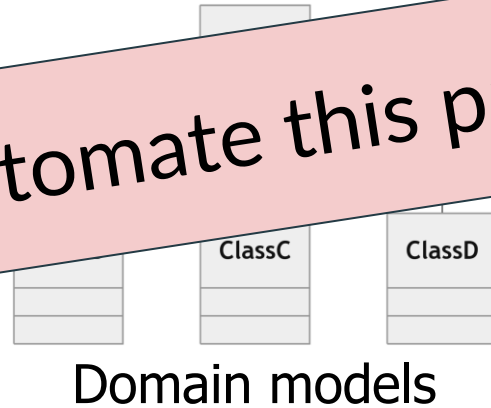
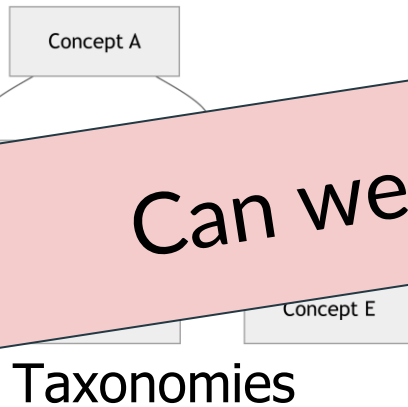


Flowcharts

Graph Models

Many graph models are used in many SE processes

Can we automate this process?



Large Language Models

Maybe we can use LLMs?

A Good Graph Model



No **syntax error**



PlantUML



Consistent

Metamodel

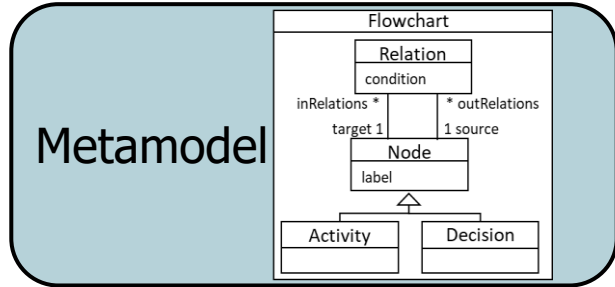
Constraint



Accurate

Textual description

LLM Generation

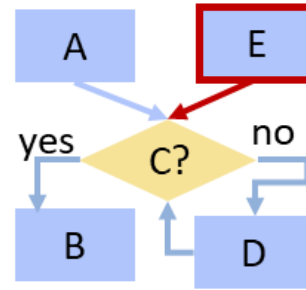


Constraints

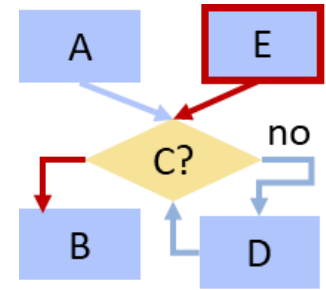
1. There is only one starting node
2. There is no self-loop
3. Decision nodes must have conditions

Textual description

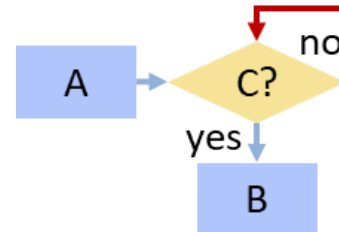
Description about activity A, B, D and condition C



Extra activity E



No condition



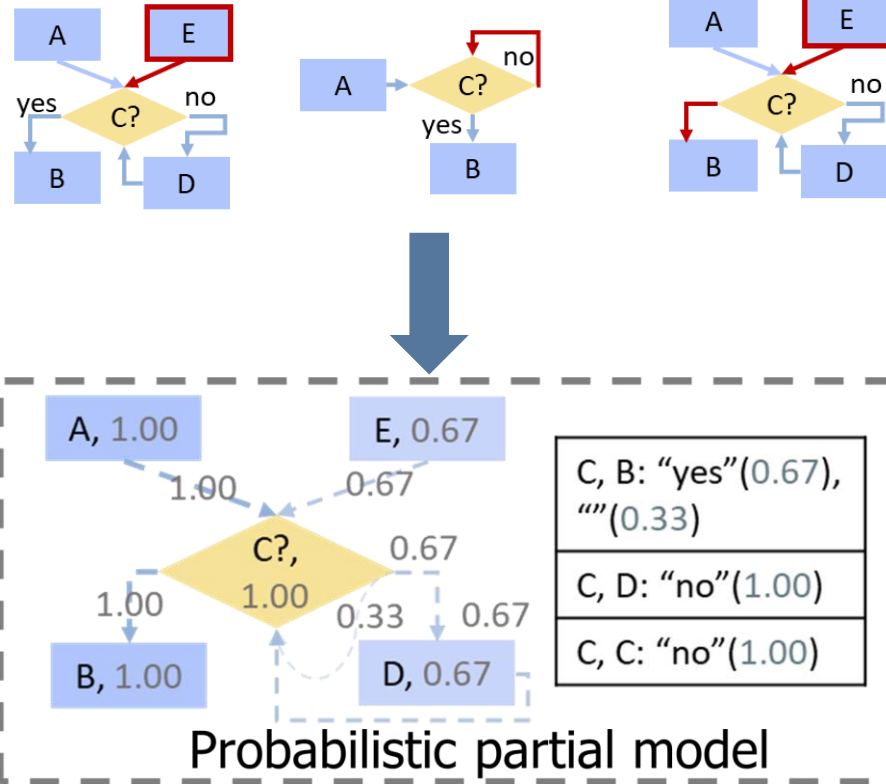
Self loop



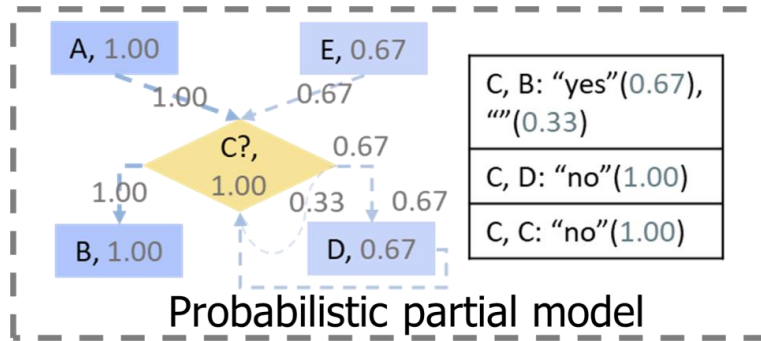
Syntax error

Abstraction

Abstraction with
node similarity and
graph matching

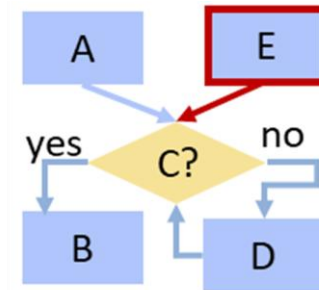


Concretization



Description about activity A, B, D and condition C

Maximize probability only
Maximum likelihood



Concretization

Constraint translation

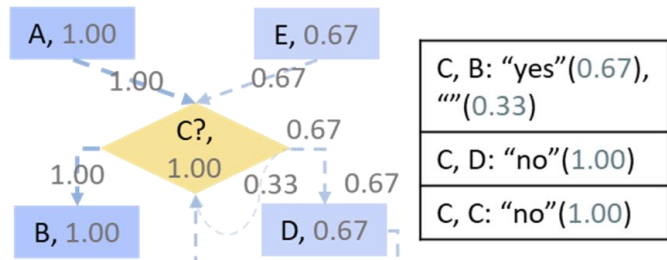
1. $\forall \text{Node } u, v; !\text{inRelations}(u, *) \wedge !\text{inRelations}(v, *) \Rightarrow u = v$
2. $\forall \text{Node } u, v; \text{Next}(u, v) \Rightarrow u \neq v$
3. $\forall \text{Relation } r, \text{Decision } d; \text{outRelations}(d, r) \Rightarrow r.\text{condition} \neq ""$

Problem formulation

Optimization

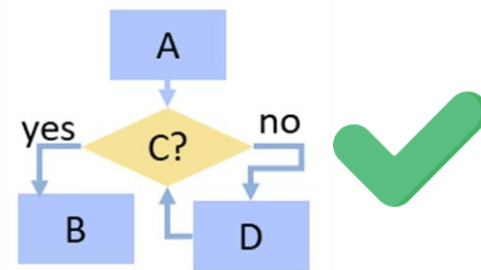


GUROBI
OPTIMIZATION

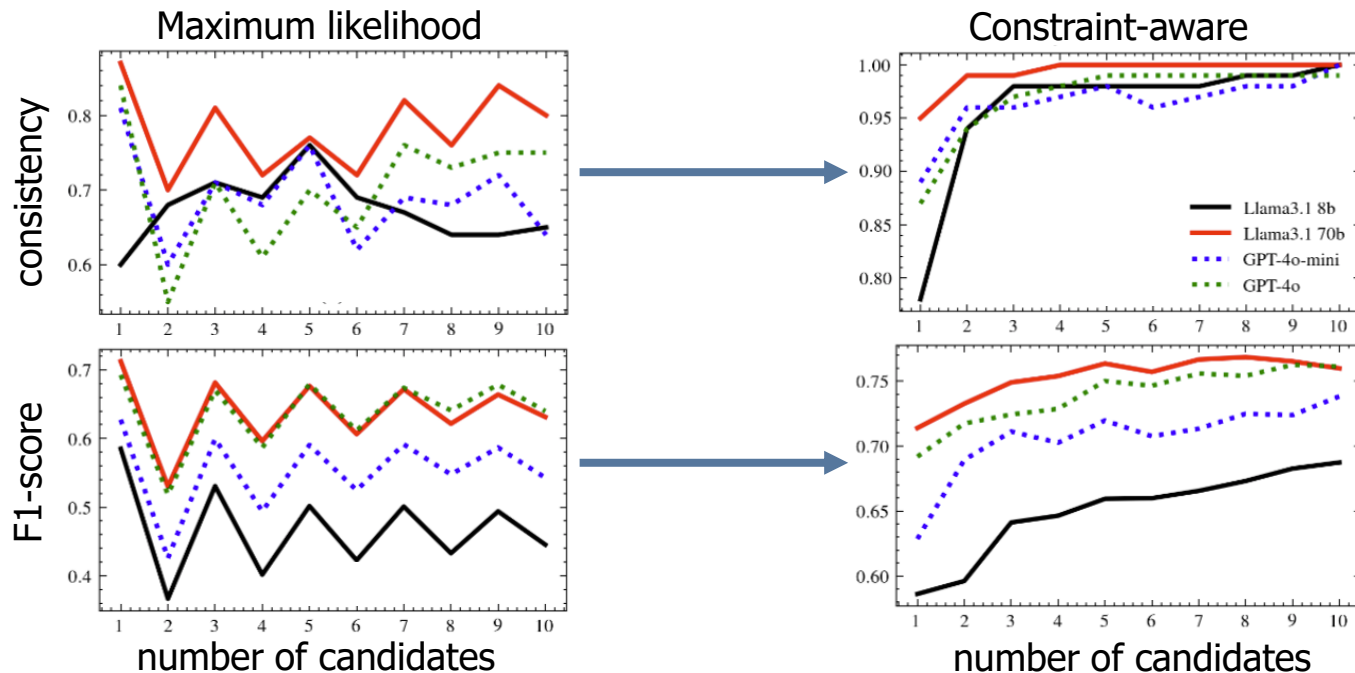


Probabilistic partial model

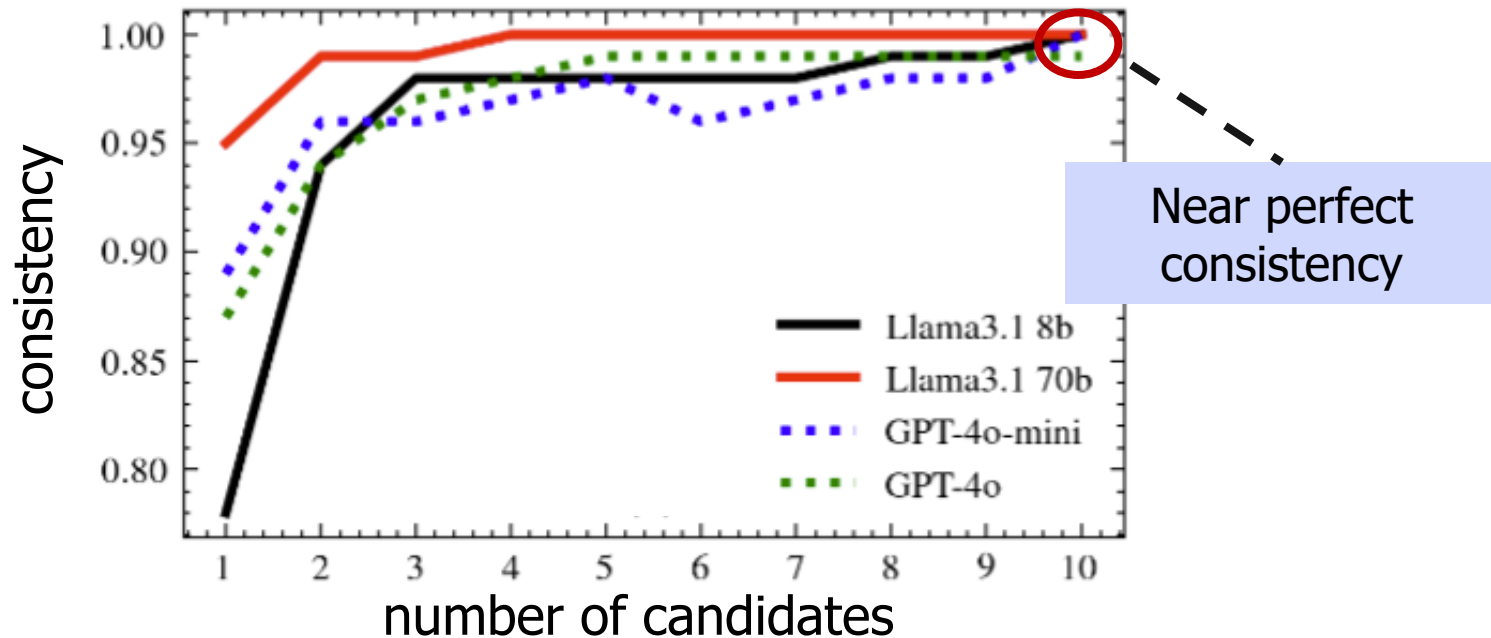
Constraint optimization



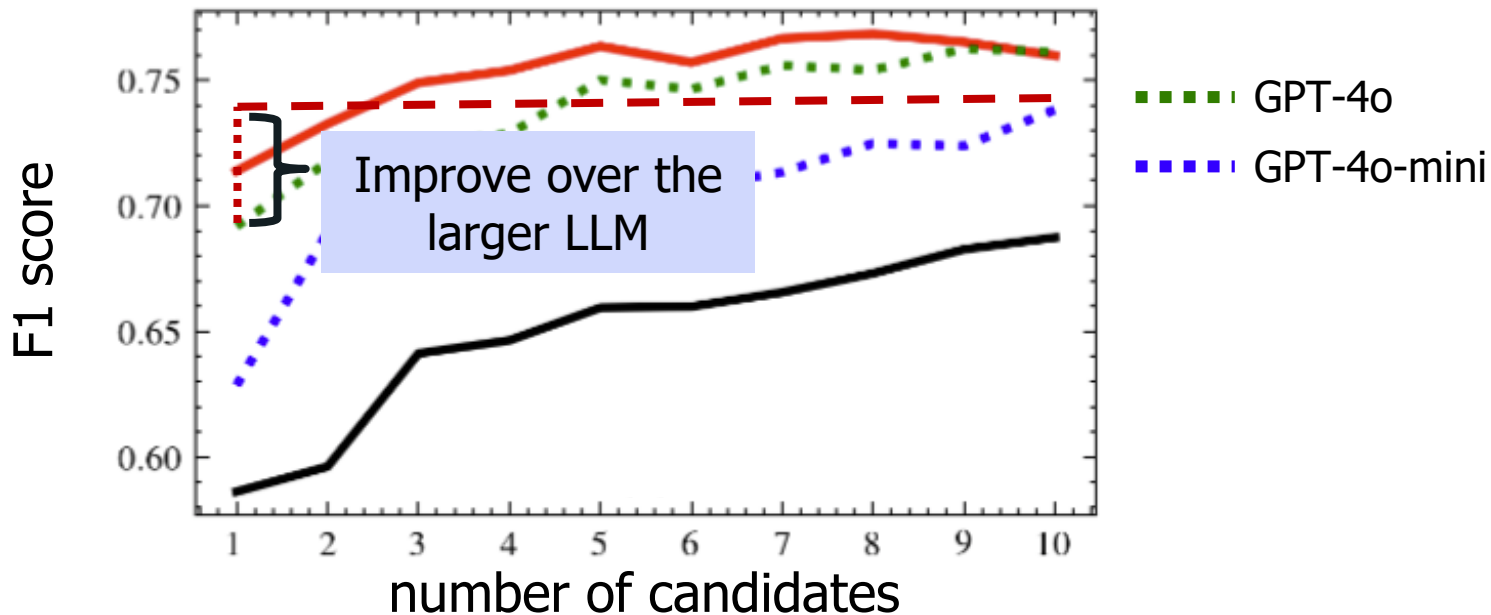
Evaluation on Taxonomies



Evaluation



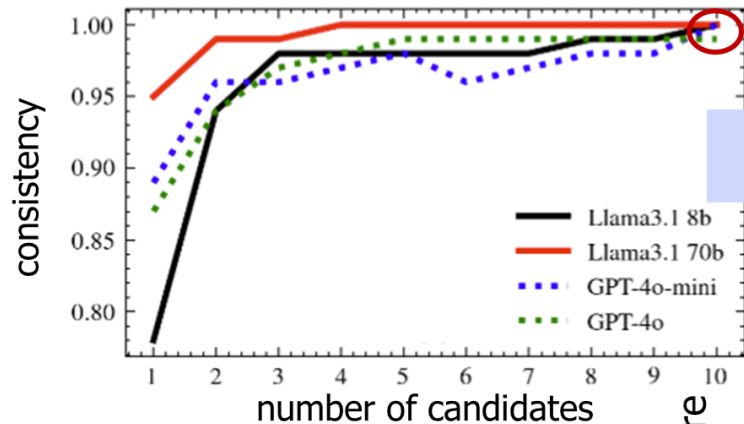
Evaluation



Uncertainty



Quality



Near perfect consistency

