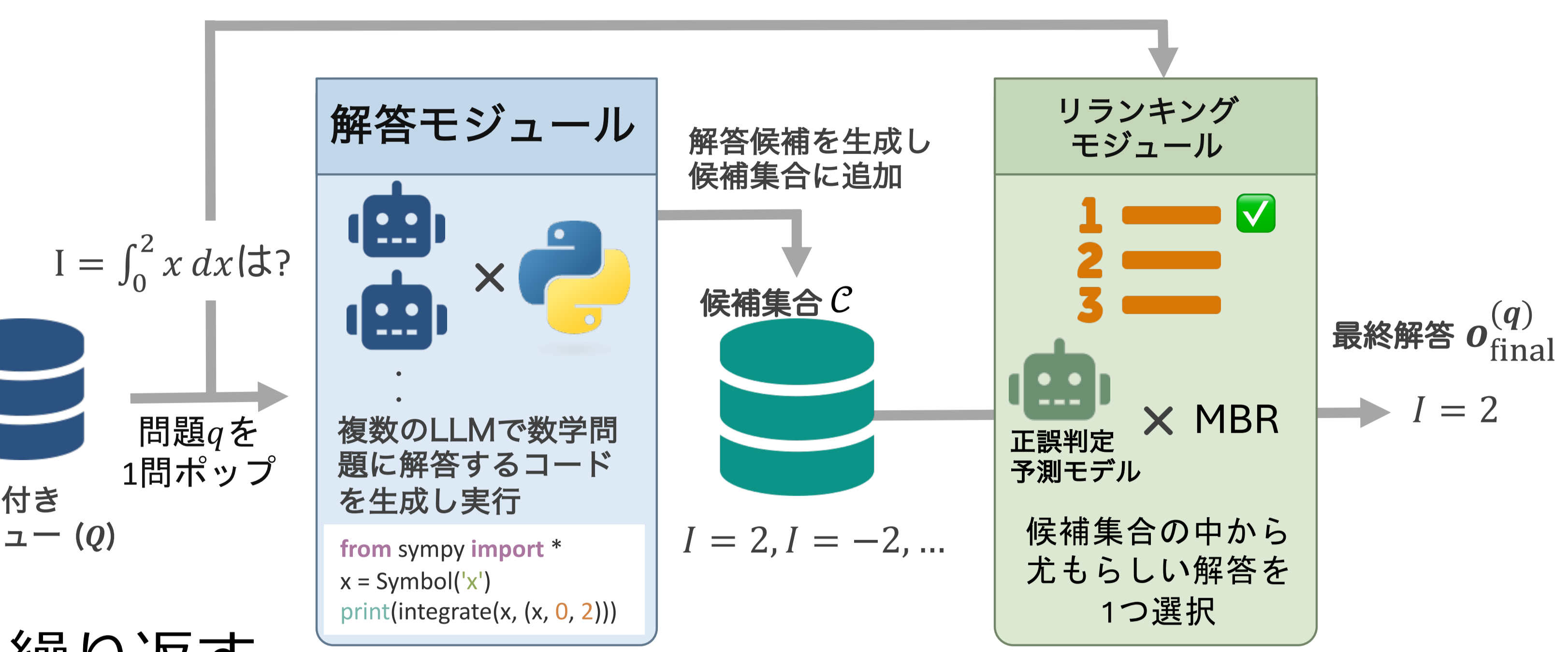


目標: LLM-jp-v4 8Bによるgpt-oss-120bの思考過程も含めた出力の模倣 + α

システム概要 / 基本戦略

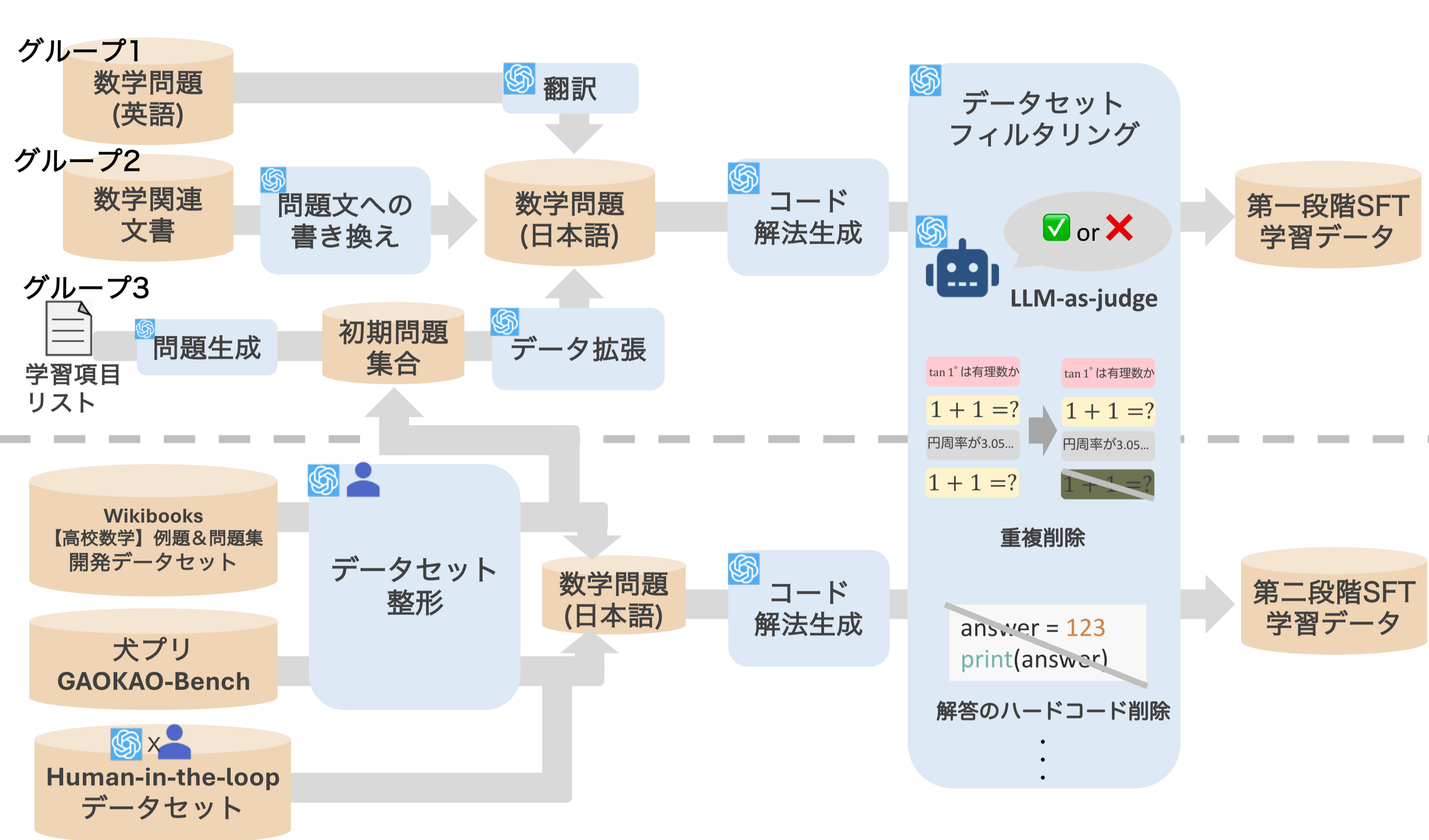
- 解答モジュール**
問題文を受け取り解答を計算するPythonコードを複数モデルで生成
- リランキングモジュール**
解答候補 (コード, 実行結果) を受け取り, 尤もらしい解答を1つ選択



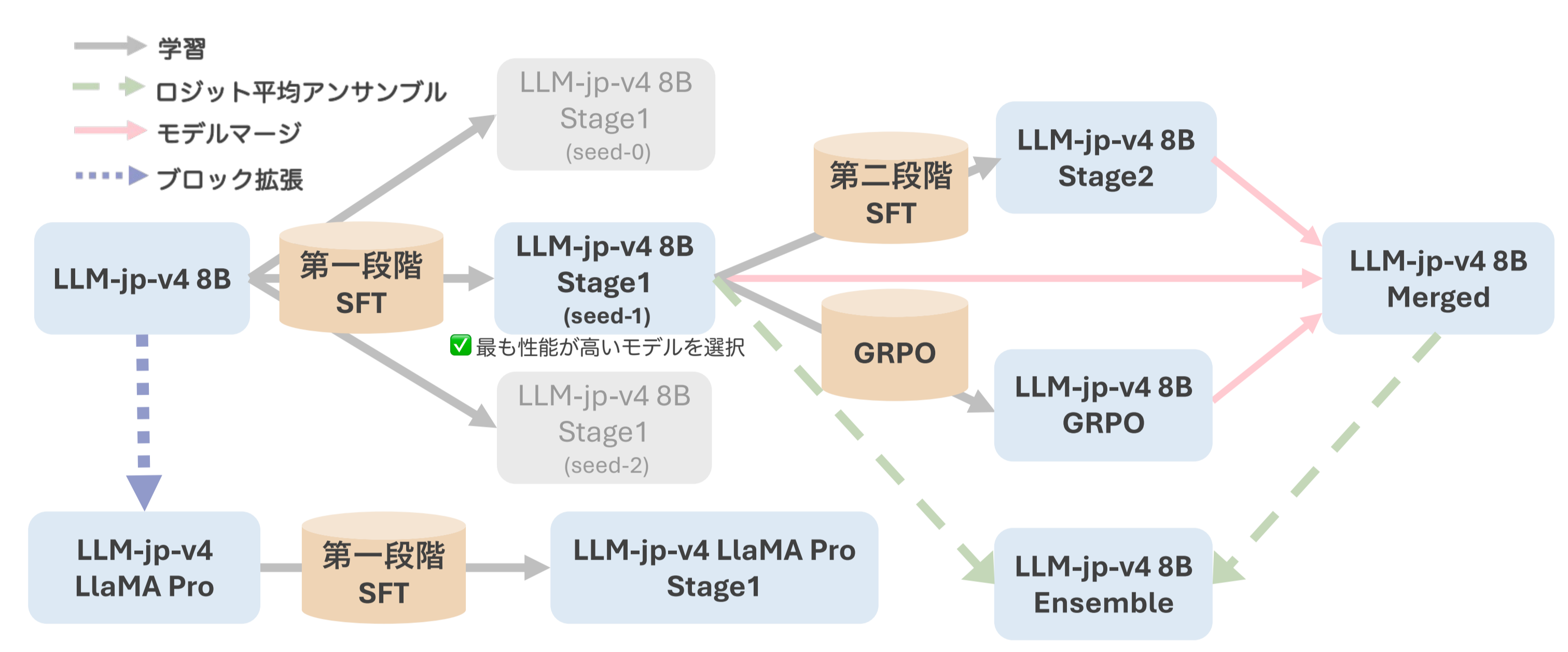
解答生成/リランキングの処理を制限時間の許す限り繰り返す

解答モジュール

データセット構築



モデル構築



- LLaMA Pro** [Wu et al. LLaMA Pro: Progressive LLaMA with Block Expansion. 2024]
同一層の繰り返しによってパラメータ数を増やす
- ロジット平均アンサンブル**
複数モデルの出力確率分布の平均を元に次のトークンを決定

リランキングモジュール

MBR (最小ベイズ複合) によるリランキングを正誤判定予測モデルで補正

正誤判定予測モデル



MBRリランキング

math-evalを用いて候補集合Cをリランキング

$$s_i = \sum_{j \neq i} \text{MathEval}(o_i, o_j).$$

- s_i : i 番目の候補のスコア
- $o_i, o_j \in C$: i, j 番目のコードの実行結果

事後評価結果

提出システム: b., c., e. のモデルを合計7モデルを組み合わせるリランキング

	math-eval \uparrow	
a. gpt-oss-120b	92.57	gpt-oss-120bと同等以上の性能
b. LLM-jp-v4 8B (Stage1)	86.75	
c. LLM-jp-v4 8B (Merged)	92.57	
d. LLM-jp-v4 8B (Merged) + Reranking	93.98	モデルの組み合わせ方に改善の余地あり
e. LLM-jp-v4 LLaMA Pro (Stage1)	89.96	
f. LLM-jp-v4 8B (Ensemble)	90.96	
g. 提出システム	91.16	
h. LLM-jp-v4 8B (Stage1, w/o thinking)	78.31	

システムの限界

Sympyに依存しているため数学的には等価だが不自然な解答を出力する可能性がある

問題: 複素数平面上の点 $z = 1 + 2i$ を原点を中心に反時計回りに $\pi/6$ 回転させた点を求めなさい。

生成されたコード:

```
from sympy import I, exp, latex, pi, simplify
z = 1 + 2 * I
theta = pi / 6
z_rot = simplify(z * exp(I * theta))
latex_answer = latex(z_rot)
print(f"Answer: {latex_answer}")
```

出力: $\sqrt{-1} (1 + 2i)$