# Jolly-MILP: A Joint LLM–Evolution Framework for MILP Solvers

*From Walls to Synergy — jointly optimizing presolving, cut selection & branching variable selection*

Ziao Guo[1] , Yuan Feng[1] , Chenhao Ying[1] , Junchi Yan[1 2] | [1] Shanghai Jiao Tong University [2] Shanghai Innovation Institute

SHANGHAI JIAO TONG UNIVERSITY

RETHINKLAB SJTU

ReThinkLab@SJTU

NEURAL INFORMATION PROCESSING SYSTEMS

## Background & Motivation

**Modern MILP** solvers (SCIP, Gurobi) run a **multi-component** pipeline — presolving, cut selection, branching. ML speeds up individual parts, but tunes each component **in isolation.**

### The Wall

Isolated components **ignore cross-component interactions**; naively stacking them yields **diminishing returns.**

### Key Idea

A single **LLM-guided evolutionary framework** co-evolves all three components jointly, explicitly modeling their **synergy.**

### Three optimized components

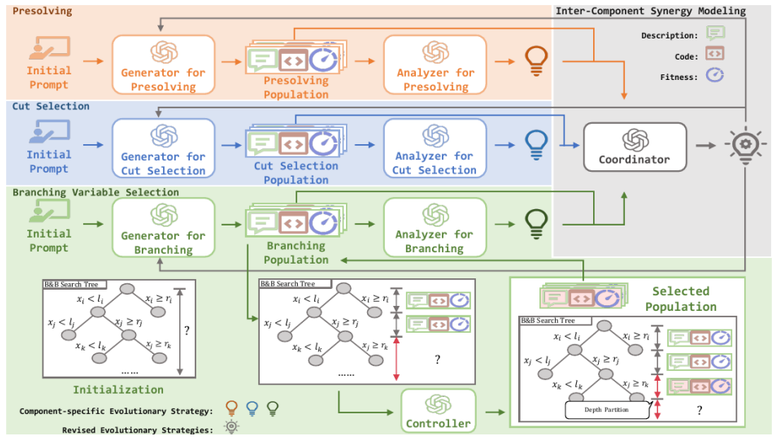
- **Presolving —** tighten & simplify the formulation before search.
- **Cut selection —** choose which cutting planes to add each round.
- **Branching var. —** pick the variable to branch on in the B&B tree.

## Contributions

1. **Jolly-MILP** — a unified LLM framework jointly optimizing presolving, cut selection & branching.
2. **Hierarchical control** — a Controller depth-partitions the B&B tree into adaptive, multi-stage branching strategies.
3. **Decoupling finding** — component- vs. system-level gains diverge: improvement comes from coordination, not any single part.

| 3 | 4 | 9/9 |
|---|---|---|
| components optimized jointly | collaborating LLM agents | datasets — best result |

## The Jolly-MILP Framework




*Per component: a Generator + Analyzer evolve symbolic strategies; a shared Coordinator aligns their co-evolution; a Controller depth-partitions the B&B tree for multi-stage branching.*

### Generator

Writes symbolic algorithms (code + rationale) per component.

### Analyzer

Diagnoses fitness gaps → proposes evolutionary directions.

### Coordinator

Models inter-component synergy → one global strategy.

### Controller

Depth-partitions tree → adaptive multi-stage branching.

### Fitness evaluation & evolutionary iteration

A **test template** fixes two components and swaps one, so joint fitness cleanly isolates each component's contribution.

Generate → **Joint-eval** → Analyze → Coordinate → Update

### Hierarchical branching control

One branching rule rarely fits the whole search (cf. strong→pseudocost). A dedicated **Controller** evolves depth-dependent strategies in stages:

- **Monitor** reads depth, solving time & node count at each depth level.
- **Partition** picks a depth point d⋆; fixes the best strategy on $[d_1, d\star)$.
- **Re-evolve** launches a fresh search for the deeper range $[d\star, d_2)$.

## Experiments & Results

### Main results — 9 datasets vs. ML & LLM baselines

| | Easy: Multiple Knapsack (n = 720, m = 72) | | | Easy: Maximum Independent Set (n = 500, m = 1953) | | | Easy: Set Covering (n=1000, m=500) | | |
|---|---|---|---|---|---|---|---|---|---|
| **Method** | Time(s) ↓ | Improv. ↑ | Effect. ↑ | Time(s) ↓ | Improv. ↑ | Effect. ↑ | Time(s) ↓ | Improv. ↑ | Effect. ↑ |
| Default | 2.57 | - | - | 13.30 | - | - | 16.47 | - | - |
| SMAC3 | 2.02 | 21.40% | 27% | 12.19 | 8.35% | 20% | 12.78 | 22.40% | 26% |
| EoH | 2.58 | -0.39% | 6% | 14.76 | -10.98% | 8% | 15.94 | 3.22% | 6% |
| ReEvo | 2.06 | 19.84% | 22% | 10.28 | 22.71% | 26% | 13.66 | 17.06% | 28% |
| Integrated-ML (config1) | 1.86 | 27.63% | 31% | 9.48 | 28.72% | 33% | 11.07 | 32.79% | 37% |
| Integrated-ML (config2) | 1.94 | 24.51% | 32% | 9.13 | 31.35% | 35% | 10.64 | 35.40% | 35% |
| **Jolly-MILP (Ours)** | **1.71** | **33.46%** | **45%** | **8.82** | **33.68%** | **48%** | **10.59** | **35.70%** | **46%** |

| | Medium: CORLAT (n = 466, m = 486) | | | Medium: MIK (n = 413, m = 346) | | | Hard: MIRP (n = 4120, m = 6857) | | |
|---|---|---|---|---|---|---|---|---|---|
| **Method** | Time(s) ↓ | Improv. ↑ | Effect. ↑ | Time(s) ↓ | Improv. ↑ | Effect. ↑ | PD Integral ↓ | Improv. ↑ | Effect. ↑ |
| Default | 27.74 | - | - | 255.20 | - | - | 4804.68 | - | - |
| SMAC3 | 22.16 | 20.12% | 36% | 245.29 | 3.88% | 10% | 4472.31 | 6.92% | 30% |
| EoH | 24.68 | 11.03% | 12% | 251.48 | 1.46% | 5% | 4727.19 | 1.61% | 10% |
| ReEvo | 23.72 | 14.49% | 20% | 248.45 | 2.64% | 10% | 4615.33 | 3.94% | 10% |
| Integrated-ML (config1) | 20.18 | 27.25% | 52% | 247.60 | 2.98% | 10% | 4326.11 | 9.96% | 25% |
| Integrated-ML (config2) | 19.22 | 30.71% | 50% | 241.15 | 5.51% | 10% | 4210.86 | 12.36% | 30% |
| **Jolly-MILP (Ours)** | **18.31** | **33.99%** | **62%** | **236.81** | **7.21%** | **15%** | **3932.75** | **18.15%** | **40%** |

| | Hard: Item Placement (n = 1083, m = 195) | | | Hard: Load Balancing (n = 61000, m = 64304) | | | Hard: Anonymous (n = 37881, m = 49603) | | |
|---|---|---|---|---|---|---|---|---|---|
| **Method** | PD Integral ↓ | Improv. ↑ | Effect. ↑ | PD Integral ↓ | Improv. ↑ | Effect. ↑ | PD Integral ↓ | Improv. ↑ | Effect. ↑ |
| Default | 304277.69 | - | - | 9462.41 | - | - | 146083.21 | - | - |
| SMAC3 | 293248.72 | 3.62% | 37% | 8957.42 | 5.34% | 34% | 117962.47 | 19.25% | 50% |
| EoH | 310867.55 | -2.17% | 11% | 9340.66 | 1.29% | 16% | 147231.44 | -0.79% | 5% |
| ReEvo | 297246.38 | 2.31% | 24% | 9153.34 | 3.27% | 27% | 121686.85 | 16.70% | 20% |
| Integrated-ML (config1) | 291340.34 | 4.25% | 46% | 8774.50 | 7.27% | 48% | 110855.60 | 24.11% | 60% |
| Integrated-ML (config2) | 289463.98 | 4.87% | 48% | 8642.29 | 8.67% | 46% | 109492.37 | 25.05% | 60% |
| **Jolly-MILP (Ours)** | **270763.40** | **11.01%** | **57%** | **8513.75** | **10.03%** | **59%** | **107324.21** | **26.53%** | **65%** |

### Generalization across problem scales (time ↓)

| | Maximum Independent Set | | Set Covering | |
|---|---|---|---|---|
| **Method** | Medium (n=1000, m=3946) | Large (n=1500, m=5940) | Medium (n=1000, m=1000) | Large (n=1000, m=2000) |
| Default | 258.71 | 2323.72 | 103.21 | 2046.75 |
| ReEvo | 236.43 | 2281.29 | 92.60 | 2087.84 |
| Integrated-ML (config2) | 202.68 | 2107.35 | 76.34 | 1719.06 |
| **Jolly-MILP (Ours)** | **190.36** | **2058.63** | **67.87** | **1703.12** |

### Ablation — Coordinator & Controller both matter (time ↓)

| **Dataset** | Full | w/o Coordinator | w/o Controller | w/o Coordinator & Controller |
|---|---|---|---|---|
| Multiple Knapsack | 1.71 | 2.13 | 2.07 | 2.50 |
| Maximum Independent Set | 8.82 | 10.15 | 9.54 | 11.26 |
| Set Covering | 10.59 | 13.24 | 11.98 | 14.29 |

### Transfer

| Train \ Evaluate | MK | MIS | SC |
|---|---|---|---|
| MK | – | 6.17% | 8.32% |
| MIS | 10.25% | – | 21.09% |
| SC | 11.13% | 19.62% | – |

### Takeaway

Explicitly modeling **cross-component synergy** — not isolated enhancement — is the key to further gains, pointing toward solver-level **co-design.**