

IB-Net: Initial Branch Network for Variable Decision in Boolean Satisfiability

Tsz Ho Chan¹, Wenyi Xiao¹, Junhua Huang¹ and Huiling Zhen¹ and Guangji Tian² and Mingxuan Yuan^{1*}

¹Noah’s Ark Lab, Huawei

²Hisilicon, Huawei

chantszho1@huawei.com, xiaowenyi2@huawei.com, huangjunhua15@huawei.com, zhenhuiling2@huawei.com, tianguangji@hisilicon.com, Yuan.Mingxuan@huawei.com

Abstract

Boolean Satisfiability problems are vital components in Electronic Design Automation, particularly within the Logic Equivalence Checking process. Currently, SAT solvers are employed for these problems and neural network is tried as assistance to solvers. However, as SAT problems in the LEC context are distinctive due to their predominantly unsatisfiability nature and a substantial proportion of UNSAT-core¹ variables, existing neural network assistance has proven unsuccessful in this specialized domain. To tackle this challenge, we propose IB-Net, an innovative framework utilizing graph neural networks and novel graph encoding techniques to model unsatisfiable problems and interact with state-of-the-art solvers. Extensive evaluations across solvers and datasets demonstrate IB-Net’s acceleration, achieving an average runtime speedup of 5.0% on industrial data and 8.3% on SAT competition data empirically. This breakthrough advances efficient solving in LEC workflows.²

1 Introduction

In computational theory and complexity, the Boolean Satisfiability (SAT) problem stands as a cornerstone. This decision problem, underpinned by the task of identifying the possibility of truth assignment to the Boolean expression, is known to be NP-complete and computationally demanding [Cook, 1971]. SAT is fundamental to various practical applications, including software verification, hardware design, and more [Kasi and Sarma, 2013] [Leino, 2010]. Those problems will first be reduced into the SAT formula and then apply SAT solvers. During the development of SAT solvers, various techniques have been applied to boost the efficiency of the solving process. There are currently two major categories, Stochastic Local Search (SLS) based solvers and Conflict Driven Clause Learning (CDCL) based solvers. SLS

solvers perform initial assignments for all variables in the formula and repeatedly flip the variable assignment to maximize the internal score, which leads to a solution. The foundation of CDCL solvers are clause learning and deep backtracking search, which make assignment for one variable each time until the conflict occurs to perform analysis and backtracking.

Following the ubiquity of SAT, its implication is particularly critical in Electronic Design Automation (EDA), the suite of software used in designing electronic systems and chips. Specifically, Logic Equivalence Checking (LEC), an essential step in verifying the correctness of a circuit design after transformation in the EDA process, leverages the efficiency of SAT solvers. LEC aims to determine whether two circuits are functionally equivalent by encoding the checking problem into Boolean expressions and applying the SAT solver. As this process is to detect the potential fault in transformation, which should exist with a small possibility, we would expect that most SAT problems in LEC are unsatisfiability (UNSAT): no possible variables assignment can be found. The SAT solver will be called by LEC for numerous times for each circuit design during the design verification stage. Thus the performance of applied SAT solvers would be the critical point to boost the efficiency of LEC process.

In recent years, there have been some studies utilizing neural networks (NN), mostly graph neural networks (GNN), to solve SAT problems directly or assist current SAT solvers. NeuroSAT pioneers the usage of NN in SAT by proposing an end-to-end model to predict the satisfiability of an SAT problem and further predict the assignment of variables [Selsam *et al.*, 2019]. NLocalSAT adopts the model design of NeuroSAT and uses the model to boost SLS solvers offline, modifying the initial variable assignment with the output of NN for one time [Zhang *et al.*, 2021]. However, it utilizes SAT/UNSAT for the instances as the supervision signal, which is invalid in the LEC verification scenario. Although Neurocore focuses on UNSAT problems, it computes the variable assignment periodically, which is not suitable for high-frequency verification in LEC [Selsam and Bjørner, 2019].

Therefore, we propose the Initial Branch Network (IB-Net), a framework that targets UNSAT problems and interacts with state-of-the-art CDCL solver to perform offline one-time branch initialization and boost efficiency of end-to-end SAT-solving process. We design Weighted Literal-Incidence

*Corresponding author

¹An UNSAT-core is a unsatisfiable subset of the original clauses in a Boolean expression. Its removal would render the problem satisfiable, aiding in identifying the cause of unsatisfiability.

²The implementation code will be provided after submission.

Graph to encode the Boolean formula and employ GNN to predict the possibility of UNSAT-core variables. By proposing novel graph node feature embedding and loss function, we address the imbalanced data distribution of problems in LEC (mostly UNSAT problems and large UNSAT-core). We evaluate IB-Net and previous works with CDCL solvers on both the open SAT Competition dataset and the real industrial dataset. IB-Net achieves a 5.0% runtime reduction per problem in the whole LEC pipeline, while other benchmarks have no effect or even runtime increase. For the open SAT Competition dataset, IB-Net gains an 8.3% runtime reduction, much higher than other works. Such experimental results show the efficiency-boosting of IB-Net. All runtimes are computed with network construction and inference time inclusive.

2 Related Work

The EDA field has leveraged the power of NN for various applications. The complexity of EDA tasks, including verification like LEC, provides a fertile ground for applying machine learning techniques to improve efficiency and robustness. In early work, these approaches focused mainly on data processing and process optimization. NN-based data argumentation and instance preprocessing are applied to accelerate the verification process [Huang *et al.*,]. Huang et al. further demonstrated the utility of NNs in EDA by designing a machine learning model to predict cut ranking for dealing with combinatorial optimization problems [Huang *et al.*, 2022].

As SAT solver is widely applied in EDA, the combination of NN and SAT problem has also been tried. Graph Neural Network has emerged as a powerful tool for capturing the relational information inherent in graph structures, which has naturally led to their application in SAT constraint problem-solving [Lopera *et al.*, 2021]. The application of GNN is approached in two ways: end-to-end solving and heuristic solver guidance. In the first vein, GNN is trained to learn the underlying structure and patterns in SAT problems and to directly predict satisfying assignments or indicate unsatisfiability [Selsam *et al.*, 2019] [Amizadeh *et al.*, 2019] [Xu *et al.*, 2018]. While the idea of using GNN as a solver is enticing, it also presents challenges. The complex nature of SAT problems leads to unreliability and low performance in models, especially in the industrial context. On the other hand, the use of GNN as heuristics to SAT solvers constitutes another effective application. In these approaches, GNN is typically employed to predict key aspects of the solving process, such as variable importance or ranking, thereby guiding the solver to solve efficiently [Selsam and Bjørner, 2019] [Zhang *et al.*, 2021] [Li *et al.*, 2018] [Balunovic *et al.*, 2018]. Unlike the first approach, these methods just provide information that potentially accelerates the solving process while the solving process is still done by the solver. Thus the reliability and complexity of SAT problem solving is ensured. In light of existing methodologies, our work IB-Net, a model designed to enhance the robustness and efficiency of SAT solvers through heuristic guidance and explore solver-GNN interaction, aims to address SAT solver usage in specific industrial scenarios.

3 IB-Net approach

3.1 Overview

To address our goal to assist the CDCL solvers targeting UNSAT problems, we propose the framework in Fig. 1. It is a combination of a neural network and a method to interact with solvers. For an input Boolean formula, it will firstly be converted into a Weighted Literal-Incidence Graph (WLIG) and fed into GNN for node feature updating, then gather the output as possibility to be UNSAT-core for variables. These output are assigned to SAT CDCL solver to initialize variable decision queue and decision score to guide the solving process. To make IB-Net more suitable for dealing with UNSAT problem, we made several improvement along the pipeline, which will be discussed in following subsections.

3.2 Graph Formulation

A propositional logic formula is a Boolean expression that contains conjunctions, disjunctions, variables, negations, and constants. The Boolean formula can be expressed in conjunctive normal form (CNF) with conjunction of disjunctions of variables (possibly negated variables) by transformation with linear length in linear time [Tseitin, 1983]. The SAT instance in CNF is a conjunction of clauses while a clause is a disjunction of literals (the variables and negated variables). The SAT instances are already converted into CNF before accessed by model and solvers, which is a common step when apply solvers.

To make use of the structure information within the formula, we convert SAT instance S into WLIG G . But in WLIG, the literal set of S are nodes and the common incidence of literals are edges, i.e. (l_1, l_2) hold edge if l_1 and l_2 appear in the same clause and the weight on edge account for the number of common incidences. The adjacent matrix A of graph G will be the input to the graph model.

3.3 Neural Network Model

We use a weighted GCN (WGCN) [Zhao *et al.*, 2021] as our first network to extract features of variables and a multi-layer perceptron (MLP) to output the possibility of variables presented in UNSAT-core. The WGCN can perfectly match the design of WLIG to utilize the edge weight and enhance the message exchange.

The WGCN accepts the adjacency matrix as input to build the graph structure. Apart from that, to address features of literals, we initialize the embedding for a literal node with vector $h \in \mathbf{R}^{2d}$. The vector h is produced by concatenating the degree of node $D \in \mathbf{R}$ and the literal type of node $T \in \mathbf{R}$ vertically and then feeding to a linear transformation $L_{init}: \mathbf{R}^{1 \times 2} \rightarrow \mathbf{R}^{1 \times 2d}$ to produce sparse node embedding.

$$h_i = L_{init}(D_i \oplus T_i), H = [h_1, h_2, \dots, h_N]^T \quad (1)$$

The node embedding vectors H will be updated by aggregating embedding from its neighbors during each iteration of WGCN. Namely, a single iteration can be represented:

$$H^{(l+1)} = \text{ReLU}(L_{out}(A' H^{(l)} \oplus \text{Flip}(H^{(l)}))), \quad (2)$$

where $H^{(l)}$ represents the node features after the l -th iteration. We detail the computation process inside the WGCN unit as follows.

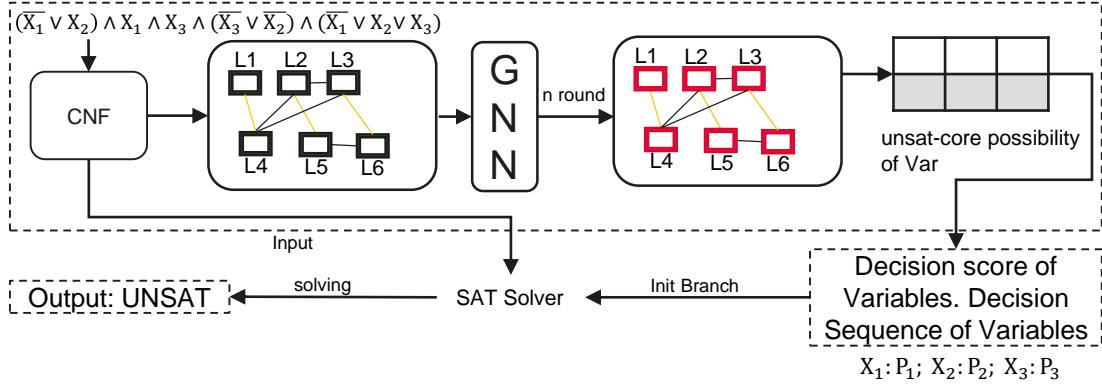


Figure 1: Model Overview

A' is the normalized adjacency matrix which encapsulates aggregation operations, and values in A' also carry weights to apply in the WGCN. A' is computed as follows to avoid gradient explosion:

$$A' = \frac{A}{\sum_{i=1}^N \sum_{j=1}^N A_{ij}}. \quad (3)$$

Function *Flip* and linear transformation $L_{out} : \mathbf{R}^{1 \times 4d} \rightarrow \mathbf{R}^{1 \times 2d}$ concatenate features of literal nodes with the corresponding negated literal nodes, forming a comprehensive representation of variables.

$$\text{Flip}(H) = [h_{\frac{N}{2}+1}, \dots, h_N, h_1, \dots, h_{\frac{N}{2}}]^T \quad (4)$$

After L iterations of WGCN, we obtain the vectors representing structural information of literals, and feed it to MLP for the UNSAT-core possibility of each variable:

$$p_i = \text{softmax}(L_2 \cdot \text{ReLU}(L_1 \cdot (h_i \oplus h_{N+1-i}) + b_1) + b_2) \quad (5)$$

3.4 Training

As we regard the UNSAT-core prediction as a classification task, the Cross-Entropy loss should be originally applied. However, due to the potential imbalance distribution (the large UNSAT-core in LEC UNSAT Problem), we utilize Focal loss [Lin *et al.*, 2020] as the loss function. The major difference between Focal loss and traditional Cross-Entropy loss would be the additional parameters $\alpha & \beta$ to balance the importance of the class with a small portion. For p stands for prediction possibility and y is truth, the loss is computed as:

$$FL = \sum_{i=1}^n (y_i(-\alpha(1-p_i)^\gamma \log(p_i)) + (1-y_i)(-(1-\alpha)p_i^\gamma \log(1-p_i))), \quad (6)$$

where n is the number of variables in the target CNF.

3.5 Interaction with Solver

In this section, we first introduce the CDCL process and the crucial role of variable decision within it. Then we show how our model targets these stages and interacts with the process.

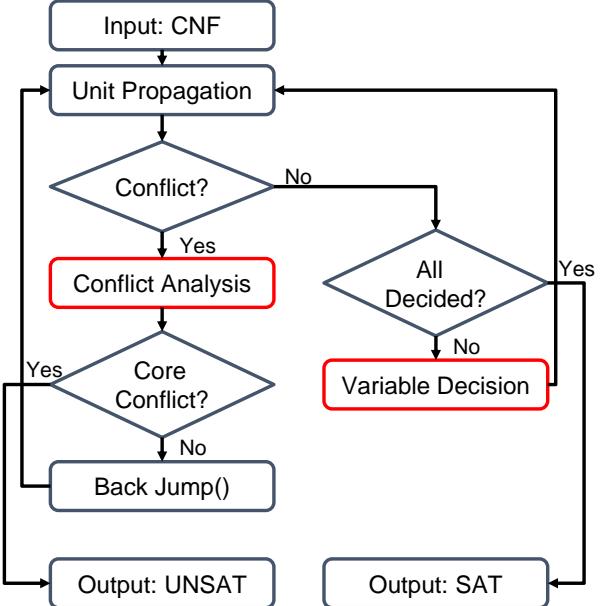


Figure 2: CDCL Process

CDCL WorkFlow

The Conflict-Driven Clause Learning algorithm, a prevailing methodology used by modern SAT solvers, functions through a combination of decision-making, unit propagation, conflict analysis, and non-chronological backtracking.

Fig. 2 provides a detailed workflow for the CDCL process.

From Fig 2, it can be found that the algorithm begins with a decision-making process where Boolean values are assigned to variables according to a decision queue. This is succeeded by unit propagation, whereby necessary assignments of other variables are deduced based on the current partial assignment. Conflicts, however, can occur if a clause in the problem becomes unsatisfiable with these assignments. In response to conflicts, the CDCL flow initiates a conflict analysis, learning and adding a new clause representing the conflict reason to the problem to avoid repetition of the same error in fu-

ture iterations. The CDCL algorithm concludes with non-chronological backtracking/backjumping, which does not merely revert the most recent decision but backjumps to a point preceding the conflict, result in assigning a lower place in the decision-making queue or a lower decision-score to these variables involved in conflict. This allows a more rapid exploration of different solution spaces and aids in the avoidance of past mistakes. Two critical stages of this process are conflict analysis (clause learning) and variable decision, which collaboratively construct the decision queue for variable branching. Ideally, a swift identification of variables within the UNSAT-core would expedite the CDCL solver’s resolution of UNSAT problems. Conversely, prematurely including variables outside UNSAT-core within the decision queue could slow down the solver. Thus we should try to guide the decision queue to optimize the performance. Specifically, we would instruct the decision queue and decision score, the internal variable used within variable decision and conflict analysis.

Queue & Score Based Branching

Against above backdrop, we propose a novel methodology for interaction with the CDCL solver. Before attempting to solve a SAT problem, our trained network undertakes an inferential analysis of the problem, thereby calculating the probability of each variable being within the UNSAT-core. These probabilities guide the solver in establishing a queue as well as a set of scores for solver branch decisions. Probabilities output are first used as the initialization for decision scores. Then they are employed to rank variables in descending queue, which is adopted as the initial value for decision queue. As predicted decision queue and decision score are based on the probability of variables within UNSAT-core, solver can deal with those variables that highly related to UNSAT-core first to address the unsatisfiability.

These decision values and decision queue, now based on our neural network inference, are imported as the solver’s initial values. The solver then can commence its operations on problem. This importing process doesn’t undermine the solver’s capacity to comprehensively address UNSAT problems, as the basic algorithm in place still centers around identifying the UNSAT-core.

The added advantage here is that our inference performs only once for each problem, thus keeping the computational burden and associated time minimal. The processing time of the whole pipeline of IB-Net (including graph formulation, inference and solver initialization) can be controlled within 1.5 second, which is regarded as a small cost comparing to the solving time. It’s worth noting that our neural network is efficient enough to run on CPUs. Consequently, our method maintains the solver’s completeness while significantly enhancing its efficiency and speed.

4 Experimental Settings

4.1 Datasets

We utilize two datasets to train and evaluate our framework and previous approaches. The first one is the industrial LEC circuit SAT instances from real-life chip design in 2023. They

| Data | | #CNF | Avg#Var | Avg#Clause |
|------|-----------|--------|---------|------------|
| LEC | Circuit 1 | 22,050 | 1,002 | 3,970 |
| | Circuit 2 | 14,840 | 952 | 3,664 |
| | Circuit 3 | 25,591 | 1,563 | 6,172 |
| | All | 62,481 | 1,220 | 4,799 |
| SAT | SAT | 1,115 | 3,065 | 74,169 |
| | UNSAT | 985 | 3,784 | 73,751 |
| | Comp | 1,171 | 3,623 | 161,140 |
| All | | 3,271 | 3,481 | 105,178 |

Table 1: Detail statistic of LEC circuit dataset and SAT Competition dataset. Noted that over 99% of instance in LEC dataset is result in UNSAT.

are extracted from the real chip development process to reflect the need for LEC process. To fit the data into GPU easily and model the real-life production constraint, we first run state-of-the-art SAT solver (Kissat [Biere and Fleury, 2022]) on these problems and filter the instance that solver can solve within 1,000 seconds (the instances that solver cannot solve within 1,000 seconds will be regarded as hard cases and sent to redesign in real-life). Apart from the industrial dataset, we also prepare the SAT Competition dataset as previous works are mostly target open datasets. We adopt the anniversary track of SAT Competition in 2022 [Balyo *et al.*, 2022]. The anniversary track is comprised of all benchmark instances used in previous SAT Competitions to ensure the coverage of problems. We also filter the anniversary track dataset in the same way with LEC dataset. The details of filtered datasets can be found in Table 1. *Halted* here means instances unsolved within given time. After preparing the datasets, we use Kissat and Drat-trim [Wetzler *et al.*, 2014] to find variable assignments for SAT problems and UNSAT-core for UNSAT problems, which serves as the supervision of our model training. When Kissat try to solve a UNSAT instance, it will produce a proof for the unsatisfiability, which will be taken by Drat-trim to produce the UNSAT-core variable. Though the UNSAT-core found by Drat-trim is not the minimal core, it still helps in solving UNSAT instances, so we will take these UNSAT-core as all the UNSAT-core referred in following sections. The supervision production just cost the same time of solving given dataset by Kissat, which can consider as easy to follow and efficient to produce.

4.2 Experimental Setup

Our experiment design commences by segregating the two previously mentioned datasets into distinct training and testing subsets, meticulously avoiding any overlaps. We allocate 80% of the data to training, reserving the remaining 20% for testing. We benchmark against several baselines: (1) The NeuroSAT [Selsam *et al.*, 2019] model based on GNN and Literal-clause graph. We follow the official implementation and used the score they set for variable as our proposed probability. (2) Modified version of NeuroSAT: NeuroCore [Selsam and Bjørner, 2019], that target UNSAT problem. We follow the official implementation and interaction periodically with solvers. (3) NLocalSAT [Zhang *et al.*, 2021] is based on NeuroSAT and interacts with SLS solver.

| Original Solvers | LEC | SAT Comp |
|------------------|------|----------|
| CaDiCaL | 350s | 209s |
| MiniSAT | 810s | 450s |
| Glucose | 529s | 308s |
| Kissat | 335s | 195s |

Table 2: Average running time of different CDCL solvers targeting random 100-sample UNSAT problems from LEC dataset and SAT Comp dataset

| Approach | LEC Circuit | | | SAT Comp | | |
|-----------|-------------|--------|--------|----------|--------|--------|
| | Acc | Pos.F1 | Neg.F1 | Acc | Pos.F1 | Neg.F1 |
| NeuroCore | 89% | 94% | 27% | 85% | 77% | 88% |
| NeuroSAT | 90% | 93% | 8% | 77% | 62% | 82% |
| NLocalSAT | 89% | 94% | 11% | 74% | 60% | 80% |
| IB-Net | 95% | 97% | 71% | 91% | 84% | 93% |

Table 3: Model UNSAT-core prediction performance on LEC dataset and SAT Competition dataset

We adopted the official implementation but changed the output to UNSAT-core prediction. Overall, we adopt the setting for each model according to their official codes or publications but change the target output to UNSAT-core prediction.

Our choice of the Kissat SAT solver, a top-performing Conflict-Driven Clause Learning based solver, as the collaborating solver, is guided by the comparative analysis with other CDCL solvers. From Table 2, the superior performance exhibited by Kissat in comparison with other CDCL peers justifies the selection for our study[Biere, 2019][Sörensson and Eén, 2005][Audemard and Simon, 2018]. As we just maintain an interaction with Kissat, instead of recompiling the Kissat solver, our work can be transferred to future top-performing solvers.

4.3 Evaluation metrics

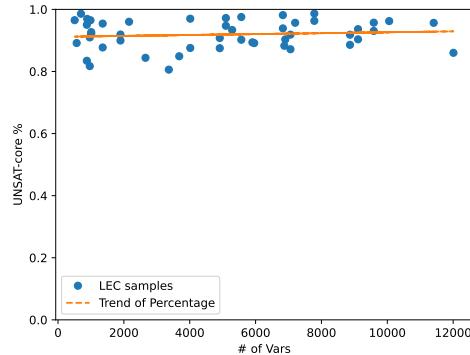
For evaluation metrics, we aim to evaluate: (1) the efficacy of UNSAT-core prediction for UNSAT problems on test datasets. (2) end-to-end time efficiency of models in interaction with Kissat solver, compared with the standalone performance of the original Kissat solver. We use evaluation metrics as: (1) A.RT means the average runtime among all CNFs. (2) Imp. means the efficiency improvement (reduction in seconds) compared with the original Kissat solver. (3) Halted represents the number of halted (time-out) CNFs within the given time (1,000 seconds). Please note that all the runtimes for NN-based solutions are network construction and inference time inclusive.

5 Experimental Results

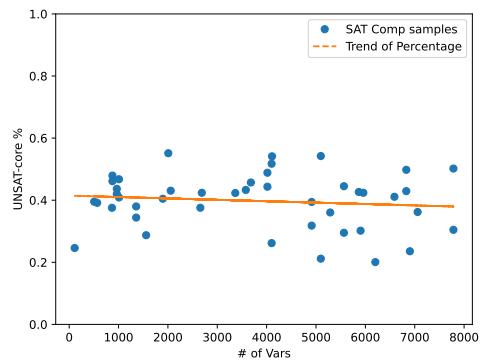
5.1 Main Results

UNSAT-core Prediction

In Table 3, we focus on the outcomes of UNSAT-core prediction. It showcases the performance of IB-Net, in comparison with other methodologies on the LEC dataset and SAT Competition dataset, which indicates that IB-Net leads the pack across both datasets. Remarkably, both F1 scores on the LEC dataset suggest IB-Net correctly identifies variables



(a) Percentage of UNSAT-core in samples with different variable sizes for LEC dataset



(b) Percentage of UNSAT-core in samples with different variable sizes for SAT Comp dataset

Figure 3: UNSAT-core Percentage

in and outside UNSAT-core. Though alternative approaches exhibit reasonable performance in SAT Competition dataset, they significantly lag behind IB-Net in LEC dataset, signifying their struggles with predicting variables in UNSAT-core accurately. This may be due to the imbalance nature of UNSAT-core in LEC UNSAT problems showed in Fig. 3: Over 90% of variables are in UNSAT-core for LEC in Fig. 3(a) while around 40% of variables are in UNSAT-core for UNSAT problems in SAT Competition in Fig. 3(b). These results underscore the efficiency of IB-Net as a potent model for UNSAT-core prediction, with its adeptness in both balance and imbalance datasets.

Runtime Reduction Performance

Table 4 provides a comparative analysis of end-to-end runtime performance for various approaches. The original Kissat is set to be the baseline. IB-Net achieves the shortest average runtime and shows improvement over the benchmark Kissat on both datasets. Specifically, IB-Net improves the average runtime by 19 seconds, accounting for 5%, and 15 seconds, accounting for 8.3%, on the LEC Circuit and SAT Competition respectively. Other methods exhibit much longer average runtime compared to Kissat and face additional halted

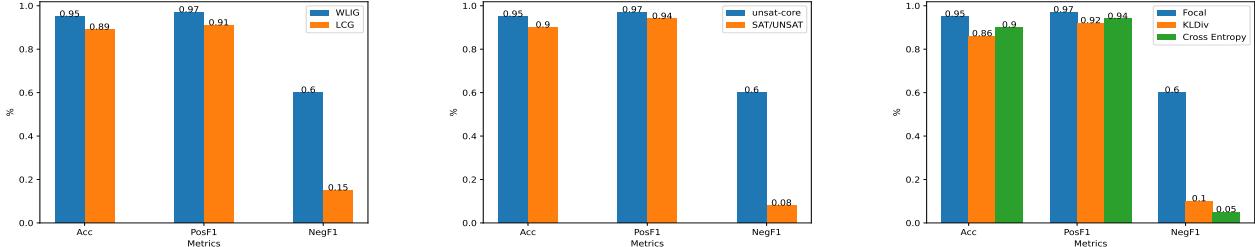


Figure 4: The ablation study

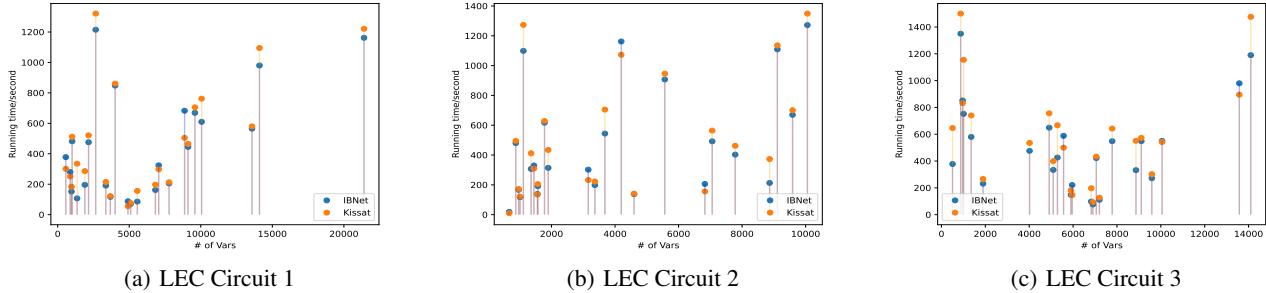


Figure 5: The circuit study

| | | Kissat | NeuroCore | NeuroSAT | NLocalSAT | IB-Net |
|----------|----------|--------|-----------|----------|-----------|--------|
| LEC | A.RT | 383s | 460s | 673s | 680s | 364s |
| | Imp. | 0s | -73s | -285s | -292s | 19s |
| | Imp. % | 0% | -18.8% | -73% | -75% | 5% |
| | Halted | 0 | 350 | 7800 | 7900 | 0 |
| | Halted % | 0% | 0.50% | 10% | 10% | 0% |
| SAT Comp | A.RT | 179s | 174s | 176s | 174s | 164s |
| | Imp. | 0s | 5s | 3s | 5s | 15s |
| | Imp. % | 0% | 3% | 2% | 3% | 8.3% |
| | Halted | 1161 | 1150 | 1161 | 1158 | 1130 |
| | Halted % | 35.5% | 35.2% | 35.5% | 35.5% | 34% |

Table 4: Model performance on LEC dataset and SAT Competition dataset

instances while IB-Net can further reduce halted instances.

As we mainly target the UNSAT problem in LEC, we present Table 5 to explore the detailed performance of different approaches on three subsets that constitute a significant proportion of the LEC industrial dataset. It can be observed that across these subsets, IB-Net consistently leads the board, even outperforming its improvements on the overall LEC data, which validates the position of IB-Net as a universally effective approach.

From Table 3 and Table 4, we can observe a consistency between the quality of UNSAT-core prediction and the improvement in runtime efficiency when dealing with UNSAT problems. It is plausible that the solver spends more time eliminating variables outside UNSAT-core from the decision queue, thus negatively impacting the performance. By fine-

| | | Kissat | NeuroCore | NeuroSAT | NLocalSAT | IB-Net |
|----------|--------|--------|-----------|----------|-----------|--------|
| Circuit1 | A.RT | 442s | 502s | 683s | 685s | 422s |
| | Imp. | 0s | -60s | -241s | -243s | 20s |
| | Imp. % | 0% | -13% | -54% | -55% | 4.5% |
| Circuit2 | A.RT | 335s | 419s | 635s | 640s | 321s |
| | Imp. | 0s | -84s | -300s | -305s | 14s |
| | Imp. % | 0% | -25% | -89% | -91% | 4.2% |
| Circuit3 | A.RT | 360s | 425s | 674s | 695s | 341s |
| | Imp. | 0s | -65s | -314s | -334s | 19s |
| | Imp. % | 0% | -18% | -87% | -92% | 5.2% |

Table 5: Model runtime performance on three main circuits in LEC dataset.

tuning the UNSAT dataset, our approach promises to offer more precise and time-efficient solution for SAT problems.

5.2 Ablation Study

We run experiments on LEC datasets with certain settings changed while other settings kept to further assess factors that may influence the effectiveness of our method on SAT problems in LEC context, shown in Fig. 4.

Graph Construction

In previous approaches, Literal-Clause Graph construction is applied as the common way to set both literal and clause as node and affiliation between clause and literal as edge. The change from LCG to WLIG give up the useless clause information in UNSAT-core prediction and focus more on literal connection.

Through experiments in Fig. 4(a), we conclude that the WLIG construction consistently outperforms Literal-Clauses Graph commonly used in previous approaches, in terms of both accuracy and F1 scores targeting UNSAT problems and especially increase the performance on variables not in UNSAT-core. This evidence supports our decision to adopt WLIG for graph construction in our model.

Supervision Signals

Results in Fig. 4(b) indicate that the UNSAT-core prediction as objective outperforms the SAT/UNSAT prediction when targeting industrial LEC dataset. The improved performance underscores the validity of our choice to employ UNSAT-core prediction as supervision signal in our model.

Loss Functions

To target the special UNSAT problem in LEC context, we amend the Cross-Entropy loss into focal loss. Upon analysis in Fig. 4(c), we find that the Focal loss delivers superior performance, outpacing both KLDiv and Cross-Entropy used in previous approaches and considered common choices in classification. This result justifies our adoption of Focal loss as the preferred loss function in the imbalance data distribution.

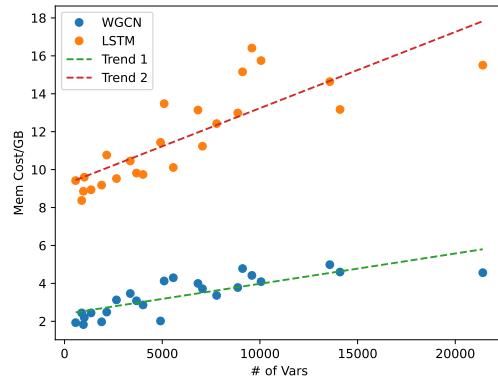
5.3 Circuit Study

To further validate the effectiveness of our proposed method, IB-Net, we conducted a comprehensive case study focusing on three circuits that constitute a significant proportion of the LEC industrial dataset. For each of these circuits, we compared the runtime of IB-Net with that of the original Kissat method. Scatter plots in Fig. 5 present the comparisons. In these plots, each point corresponds to a single SAT problem from one of the three targeted circuits.

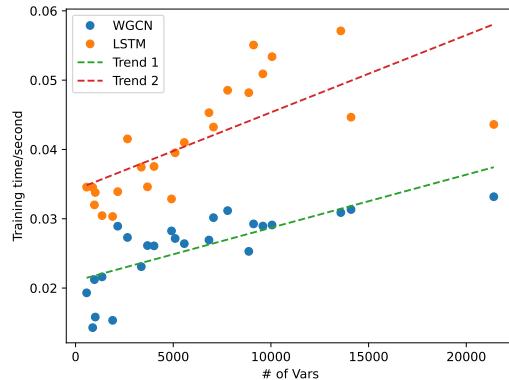
The plots illustrate a clear trend: across a wide range of problem sizes, IB-Net consistently outperforms Kissat. This performance advantage of IB-Net holds true when facing substantial variability in size of problems within these circuits, which indicate that our method is not only faster but also scalable and capable of handling large and complex problems efficiently. These findings underscore the potential of IB-Net to improve the performance of SAT solver in LEC, confirming the practical applicability of our method in a real-world setting.

5.4 Scalability

We plot scatter graphs in Fig. 6, showing memory consumption and training time of varying sizes of data samples under different computation units. WGCN uses significantly less memory across all sizes compared to LSTM, around only 30%. This stark difference suggests that model with WGCN is memory-efficient, potentially allowing for large or complex problem-solving without compromising system resources like GPU. WGCN also demonstrates faster training speeds across all data sample sizes. On average, WGCN is found to be approximately 40% quicker than LSTM. This speed advantage can reduce the time required for training models, thereby increasing the efficiency of model development process.



(a) Memory Cost of different neural network



(b) Training time of different neural network

Figure 6: The scalability: The memory cost (a) and training time (b) according to different number of variables.

6 Conclusion

In this paper, we presented a novel approach specifically tailored to the UNSAT problem solving in LEC process. Our approach, termed as IB-Net, introduces a GNN model to predict a decision strategy queue and decision scores, serving as a powerful assistant tool for CDCL SAT solvers.

Our comprehensive set of experiments established the superiority of IB-Net in accelerating SAT solving, significantly outperforming existing methods. This performance gain is particularly pronounced within the context of LEC. Furthermore, our approach has demonstrated the capability to successfully solve instances that were previously deemed unsolvable using traditional SAT solvers.

The impact of IB-Net extends beyond its immediate performance advantages. By combining machine learning techniques with traditional SAT solvers, the unique blend of these disciplines can enable more efficient and robust solvers. As we continue to refine and expand on our method, we envision a future where the performance gains offered by approaches like IB-Net become standard in the realm of SAT solving.

References

- [Amizadeh *et al.*, 2019] Saeed Amizadeh, Sergiy Matusevych, and Markus Weimer. Learning to solve circuit-sat: An unsupervised differentiable approach. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019.
- [Audemard and Simon, 2018] Gilles Audemard and Laurent Simon. On the Glucose SAT solver. *International Journal on Artificial Intelligence Tools (IJAIT)*, 27(1):1–25, 2018.
- [Balunovic *et al.*, 2018] Mislav Balunovic, Pavol Bielik, and Martin Vechev. Learning to solve smt formulas. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.
- [Balyo *et al.*, 2022] Tomas Balyo, Marijn J.H. Heule, Markus Iser, Matti Järvisalo, and Martin Suda, editors. *Proceedings of SAT Competition 2022: Solver and Benchmark Descriptions*. Department of Computer Science Series of Publications B. Department of Computer Science, University of Helsinki, Finland, 2022.
- [Biere and Fleury, 2022] Armin Biere and Mathias Fleury. Gimsatul, IsaSAT and Kissat entering the SAT Competition 2022. In Tomas Balyo, Marijn Heule, Markus Iser, Matti Järvisalo, and Martin Suda, editors, *Proc. of SAT Competition 2022 – Solver and Benchmark Descriptions*, volume B-2022-1 of *Department of Computer Science Series of Publications B*, pages 10–11. University of Helsinki, 2022.
- [Biere, 2019] Armin Biere. Cadical at the sat race 2019. 2019.
- [Cook, 1971] Stephen A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing, STOC ’71*, page 151–158, New York, NY, USA, 1971. Association for Computing Machinery.
- [Huang *et al.*,] Junhua Huang, Hui-Ling Zhen, Naixing Wang, Mingxuan Yuan, Hui Mao, Yu Huang, and Jiping Tao. Accelerate sat-based atpg via preprocessing and new conflict management heuristics. In *ASP-DAC 2022*.
- [Huang *et al.*, 2022] Zeren Huang, Kerong Wang, Furui Liu, Hui-Ling Zhen, Weinan Zhang, Mingxuan Yuan, Jianye Hao, Yong Yu, and Jun Wang. Learning to select cuts for efficient mixed-integer programming. *Pattern Recognition*, 123:108353, 2022.
- [Kasi and Sarma, 2013] Bakhtiar Khan Kasi and Anita Sarma. Cassandra: Proactive conflict minimization through optimized task scheduling. In *ICSE 2023*, pages 732–741, 2013.
- [Leino, 2010] K. Rustan M. Leino. Dafny: An automatic program verifier for functional correctness. In Edmund M. Clarke and Andrei Voronkov, editors, *Logic for Programming, Artificial Intelligence, and Reasoning*, pages 348–370, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
- [Li *et al.*, 2018] Zhuwen Li, Qifeng Chen, and Vladlen Koltun. Combinatorial optimization with graph convolutional networks and guided tree search. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems, NIPS’18*, page 537–546, Red Hook, NY, USA, 2018. Curran Associates Inc.
- [Lin *et al.*, 2020] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2):318–327, 2020.
- [Lopera *et al.*, 2021] Daniela Sánchez Lopera, Lorenzo Servadei, Gamze Naz Kiprit, Souvik Hazra, Robert Wille, and Wolfgang Ecker. A survey of graph neural networks for electronic design automation. In *2021 ACM/IEEE 3rd Workshop on Machine Learning for CAD (MLCAD)*, pages 1–6, 2021.
- [Selsam and Bjørner, 2019] Daniel Selsam and Nikolaj Bjørner. Guiding high-performance sat solvers with unsat-core predictions. In Mikoláš Janota and Inês Lynce, editors, *Theory and Applications of Satisfiability Testing – SAT 2019*, pages 336–353, Cham, 2019. Springer International Publishing.
- [Selsam *et al.*, 2019] Daniel Selsam, Matthew Lamm, Benedikt Bünz, Percy Liang, Leonardo de Moura, and David L. Dill. Learning a SAT solver from single-bit supervision. In *ICLR 2019*. OpenReview.net, 2019.
- [Sörensson and Eén, 2005] Niklas Sörensson and Niklas Eén. Minisat v1.13 - a sat solver with conflict-clause minimization. 2005.
- [Tseitin, 1983] G. S. Tseitin. *On the Complexity of Derivation in Propositional Calculus*, pages 466–483. Springer Berlin Heidelberg, Berlin, Heidelberg, 1983.
- [Wetzler *et al.*, 2014] Nathan Wetzler, Marijn J. H. Heule, and Warren A. Hunt. Drat-trim: Efficient checking and trimming using expressive clausal proofs. In Carsten Sinz and Uwe Egly, editors, *Theory and Applications of Satisfiability Testing – SAT 2014*, pages 422–429, Cham, 2014. Springer International Publishing.
- [Xu *et al.*, 2018] Hong Xu, Sven Koenig, and T. K. Satish Kumar. Towards effective deep learning for constraint satisfaction problems. In John Hooker, editor, *Principles and Practice of Constraint Programming*, pages 588–597, Cham, 2018. Springer International Publishing.
- [Zhang *et al.*, 2021] Wenjie Zhang, Zeyu Sun, Qihao Zhu, Ge Li, Shaowei Cai, Yingfei Xiong, and Lu Zhang. Nlo-calsat: Boosting local search with solution prediction. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, 2021.
- [Zhao *et al.*, 2021] Yunxiang Zhao, Jianzhong Qi, Qingwei Liu, and Rui Zhang. Wgcn: Graph convolutional networks with weighted structural features. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR ’21*, page 624–633, New York, NY, USA, 2021. Association for Computing Machinery.