

Result Quality of 3SAT on a Quantum Annealing Platform

Thomas Gabor¹, Sebastian Zielinski¹, Sebastian Feld¹, Christoph Roch¹,
Christian Seidel², Florian Neukart³, Isabella Galter², Wolfgang Maurer⁴, and
Claudia Linnhoff-Popien¹

¹ LMU Munich

² ???

³ Volkswagen Group of America

⁴ OTH Regensburg

Abstract. The abstract should briefly summarize the contents of the paper in 150–250 words.

Keywords: Quantum Computing · Quantum Annealing · 3SAT · NP · phase transition.

1 Introduction

Following the great paper [1] we thought, let's do another one!

0.5 Introduction

2 Preliminaries

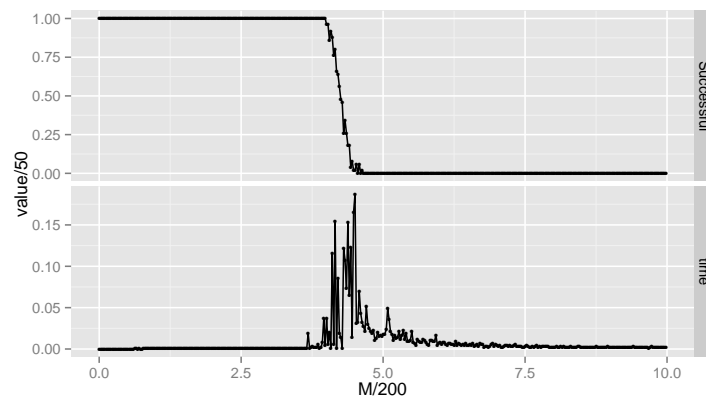


Fig.1. Lösungsfindungsdauer von 3SAT-Instanzen mit 1000 Klauseln für verschiedene Anzahl von Variablen.

Das K - Erfüllbarkeitsproblem (im Folgenden *KSAT* genannt) gehört zur Klasse der NP-Vollständigen Probleme. KSAT ist, für $K > 3$, ein zentrales Problem in der kombinatorischen Optimierung und war das erste Problem, für das die NP-Vollständigkeit gezeigt werden konnte [?].

Problembeschreibung: K - Erfüllbarkeitsproblem (nach [?]) Gegeben sei eine Menge \mathcal{B} , bestehend aus N booleschen Variablen. Aus dieser Menge \mathcal{B} werden dann K Variablen ausgewählt. Die ausgewählten Variablen, oder deren Verneinungen, werden dann durch $(K-1)$ Oder-Operatoren zu einer *Klausel* zusammengefasst. Auf diese Art und Weise werden M Klauseln gebildet. Durch anwenden von $(M-1)$ Und - Operatoren entsteht so eine KSAT-Instanz.

Das KSAT-Problem ist nun die Frage, ob eine Belegung der booleschen Variablen aus \mathcal{B} existiert, so dass eine gegebene KSAT-Instanz erfüllbar ist.

In dieser Arbeit werden ausschließlich zufällig generierte KSAT-Instanzen für $K=3$ betrachtet. Diese Probleme werden dann 3SAT genannt und sind nach [?] NP-Vollständig.

2.1 Kritischer Punkt bei zufällig generierten KSAT-Instanzen

Bei zufällig generierten KSAT-Instanzen kann beobachtet werden, die Wahrscheinlichkeit, eine korrekte Lösung für die Instanz zu finden, abrupt sinkt, wenn ein kritischer Wert α_c (gebildet aus dem Verhältnis von Anzahl der Klauseln zu Anzahl der Variablen) überschritten wird [?].

Nach [?] liegt dieser kritische Punkt bei zufällig erzeugten 3SAT-Instanzen bei

$\alpha_c \approx 4,267$. In der Umgebung des kritischen Punkts ist die Lösungsfindung (damit ist hier nicht nur eine konkrete Belegung gemeint, falls die Instanz lösbar ist, sondern auch die Erkenntnis, dass diese Instanz nicht gelöst werden kann) algorithmisch komplex. Abb. ?? verdeutlicht dieses Phänomen visuell.

Zur Erstellung von Abb. ?? wurden mit Hilfe des *Tough SAT Generators*⁵ 3SAT-Instanzen zu je 1000 Klauseln erstellt. Die Größe der Menge der booleschen Variablen wurde ausgehend von 1000 schrittweise um 10 verringert, bis ein abschließendes Klausel zu Variablenverhältnis von ca. 7 erreicht wurde. Für jede Größe der so erzeugten Variablenmenge wurden 10 Instanzen erzeugt. Die verschiedenen Instanzen wurden dann mit dem SAT-Solver *Minisat*⁶ gelöst. In Abb. ?? ist dann einmal die durchschnittliche sowie die maximale Lösungsdauer der 10 Instanzen für ein bestimmtes Verhältnis dargestellt.

2.2 Das Problem der maximal gewichteten unabhängigen Menge

–Unabhängige Menge [?] Gegeben sei ein Graph $\mathcal{G} = (V, E)$, wobei V die Menge der Knoten des Graphs ist und E die Menge der Kanten des Graphs.

Eine *unabhängige Menge* von \mathcal{G} ist eine Menge $V' \subset V$ mit folgender Eigenschaft:

$$(i, j) \notin E \quad \forall i, j \in V' \quad (1)$$

⁵ <https://toughsat.appspot.com/>

⁶ <http://minisat.se/>

Eine *unabhängige Menge* eines Graphen \mathcal{G} ist also eine Teilmenge V' der Knoten von \mathcal{G} so, dass die Knoten aus V' paarweise nicht direkt durch eine Kante von \mathcal{G} verbunden sind.

–Maximal gewichtete unabhängige Menge [?] Gegeben sei ein ungerichteter Graph $\mathcal{G} = (V, E)$, wobei V die Menge der Knoten und E die Menge der Kanten von \mathcal{G} sind, bei dem jeder Knoten $v \in V$ mit einem positiven reellen Gewicht c_v versehen wird.

Das Problem ist es dann eine Menge $S \subset V$ zu finden, so dass S eine unabhängige Menge im Sinne von Def. 2.2 ist und das Gesamtgewicht von S , gegeben durch $\sum_{v \in S} c_v$, maximal ist. Die optimale Lösung wird auch mit $WMIS(\mathcal{G})$ bezeichnet. Das Problem der maximal gewichteten unabhängigen Menge wird im Folgenden *WMIS* (engl. *Weighted Maximum Independent Set*) genannt. Für dieses Problem existiert die folgende QUBO-Darstellung:

Es sei der gleiche Graph \mathcal{G} wie aus der vorhergehenden Problembeschreibung gegeben. Die QUBO-Formulierung des WMIS-Problems lautet dann:

$$\text{Maximiere } \mathcal{Y}(x_1, \dots, x_n) = \sum_{i \in V} c_i x_i - \sum_{i,j \in E} J_{ij} x_i x_j \quad (2)$$

Ist in gerade definierter QUBO-Darstellung des WMIS-Problems $J_{ij} > \min\{c_i, c_j\}$ für alle $i, j \in E$ dann ist das WMIS von \mathcal{G} gegeben durch $WMIS(\mathcal{G}) = \{v \in V : x_v^* = 1\}$, wobei $(x_1^*, \dots, x_n^*) = \arg \max_{(x_1, \dots, x_n) \in \{0,1\}^n} \mathcal{Y}(x_1, \dots, x_n)$.

Es sei $(x_1^*, \dots, x_n^*) = \arg \max_{(x_1, \dots, x_n) \in \{0,1\}^n} \mathcal{Y}(x_1, \dots, x_n)$. Weiter sei $S^* = \{v \in V : x_v^* = 1\}$ gegeben. Angenommen, S^* ist keine unabhängige Menge. Es gibt also zwei Knoten $x, y \in S^*$ die durch eine Kante aus \mathcal{G} verbunden sind. Ohne Beschränkung der Allgemeinheit kann angenommen werden, dass $c_x < c_y$ gilt. Als nächstes entfernt man x aus S^* und bezeichnet mit $S' = S^* \setminus \{x\}$ die so entstandene Menge. Das Gewicht von S' unterscheidet sich dann durch das Gewicht von S^* durch den Ausdruck $-c_x + \sum_{j \in \text{nbr}(y) \cap S^*} J_{ij} > -c_y + J_{yz} > 0$, mit $\text{nbr}(y) = \{z \in V : (z, y) \in E\}$. Das bedeutet also S' hat ein größeres Gewicht als S^* , was einen Widerspruch zur Optimalität von S^* darstellt. Die Annahme muss also falsch gewesen sein und S^* muss eine unabhängige Menge sein.

3 Related Work

Seit über 30 Jahren ist das Lösen von NP-vollständigen bzw. NP-harten Problemen Gegenstand der Forschung (vgl. [?]). Es haben sich daher zahlreiche Methoden entwickelt um diese Probleme zu lösen. So gibt es Methoden diese Probleme zum Beispiel mittels Genetischer Algorithmen (vgl. [?], [?]), Neuronaler Netze (vgl. [?], [?]) oder Dynamischer Programmierung (vgl. [?]) zu lösen. Formuliert man ein NP-vollständiges bzw. NP-hartes Entscheidungsproblem als Optimierungsproblem, so können auch geläufige Methoden zum Lösen von Optimierungsproblemen zur Lösung herangezogen werden. Diese Methoden umfassen unter anderem die Tabu-Suche (vgl. [?], [?]) und Simulated Annealing (vgl. [?], [?]). Verschiedene NP-schwere Probleme können außerdem, trotz der Zugehörigkeit zur gleichen Komplexitätsklasse, verschieden schwierig zu lösen sein.

1 Related Work

So ist bekannt, dass manche Probleme, wie zum Beispiel das Knapsack-Problem ein FPTAS (engl. fully polynomial-time approximation scheme) besitzen. Das bedeutet, dass eine Lösung, die $(1+\epsilon)$ -nah, $\epsilon > 0$ von der optimalen Lösung entfernt ist in polynomialer Zeit, nämlich $1/\epsilon$, gefunden werden kann [?]. Über die Schwierigkeit von NP-vollständigen Problemen kann ebenfalls etwas ausgesagt werden. Es ist bekannt, dass NP-vollständige Probleme ein Phasenübergangsverhalten aufweisen können (vgl. [?], [?]). Diese Phasenübergangsgrenze teilt den Problemraum in zwei Regionen. In der einen Region kann eine Lösung relativ leicht gefunden werden, da die Lösungsdichte für diese Probleme hoch ist, in der anderen Region ist es sehr unwahrscheinlich, dass Probleme überhaupt eine korrekte Lösung enthalten können. Sehr schwer zu lösende Probleme befinden sich direkt an der Phasenübergangsgrenze [?]. Bei den in dieser Arbeit betrachteten 3SAT-Problemen liegt diese Phasenübergangsgrenze bei einem Klausel zu Variablenverhältnis von ca. 4.267 [?].

In dieser Arbeit wird zum Lösen von 3SAT-Problemen, welche NP-vollständig sind [?], das Quantum Annealing benutzt. Die Idee, die Lösung eines zu berechnenden Problems als Grundzustand eines Quantenhamiltonians zu codieren (= Quantum Annealing) Entstand ca. 1988 [?]. Ausgangspunkt dieser Art kombinatorische Optimierungsprobleme zu lösen waren Apolloni et al. (vgl. [?], [?]). In den folgenden Jahren wurde das Quantum Annealing mehrfach neu erfunden [?] (vgl. [?], [?], [?]). Seit der ersten Erfindung des Quantum Annealings wurde für viele NP-vollständige Probleme eine Problemformulierung gefunden, so dass diese Probleme auf einem Quantenannealer gelöst werden können (vgl. [?]) und bis heute findet viel Forschung für Probleme wie das Problem des Handlungsreisenden oder das 3SAT-Problem im Kontext des Quantencomputings statt (vgl. [?], [?], [?], [?], [?]). Neben dem Quantum Annealing gibt es ein weiteres Model des Quantencomputings, genannt Quantum-Gate-Computing, welches polynomiell Äquivalent zum Model des Quantum Annealings ist [?], welches in dieser Arbeit jedoch nicht berücksichtigt wird. Mit der Erfindung des Quantencomputings, einer neuen Herangehensweise an das Lösen von Problemen, entstanden auch neue Komplexitätsklassen (vgl. [?], [?]), deren Beziehungen zur klassischen Komplexitätstheorie ebenfalls Teil der Forschung war (vgl. [?], [?]).

Zu den in dieser Arbeit im Kontext des Quantencomputing betrachteten 3SAT-Problemen fand bereits viel Forschung statt. So wurden verschiedene Methoden vorgestellt, diese Probleme auf einem Quantenannealer zu lösen (vgl. [?], [?], [?]). Ebenso fanden Untersuchungen zur Lösungskomplexität von 3SAT-Problemen im Kontext des Quantencomputings statt (vgl. [?], [?]) und für bestimmte 3SAT-Instanzen die nicht effizient von einfachen Quantenalgorithmien gelöst werden können wurde eine mögliche Lösungsmethode vorgeschlagen (vgl. [?]). Ebenso wurden Variationen des 3SAT-Problems, wie zum Beispiel das Exact Cover-Problem angewandt auf 3SAT-Formeln, betrachtet (vgl. [?]).

4 Experimental Setup

4.1 Optimierungs-Postprocessing des D-Wave

Nachdem ein Problem via Quantumannealing gelöst wurde, stellt der D-Wave Quantenannealer zusätzlich die Möglichkeit verschiedener Arten von Postprocessing bereit. In dieser Arbeit wird ausschließlich das Optimierungs-Postprocessing des D-Wave benutzt.

Das Ziel des Optimierungs-Postprocessings ist es, eine Menge an Lösungen zu erhalten, die auf einem bestimmten Graphen \mathcal{G} die geringste Energie aufweisen. Um dies zu erreichen, wird ein gegebener Graph \mathcal{G} durch eine heuristische Methode (vgl. [?]) in mehrere Teilgraphen zerlegt. Danach wird die bereits von der QPU erhaltene Lösung so verändert, dass für jeden Subgraphen eine lokal optimale Lösung entsteht. Für eine tiefergehende Darstellung des Postprocessings auf dem D-Wave sei an dieser Stelle auf die D-Wave Dokumentation zu diesem Thema (siehe [?]) verwiesen, in welcher auch der hier dargestellte grobe Überblick über das Optimierungs-Postprocessing gefunden werden kann.

4.2 Logisches Postprocessing

Zusätzlich zum D-Wave Optimierungs-Postprocessing wird in dieser Arbeit ein vom Autor dieser Arbeit erdachtes logisches Postprocessing verwendet, um die ursprünglichen Ergebnisse der QPU weiter zu verbessern. Bei diesem logischen Postprocessing, wird eine von der QPU erhaltene Lösung, welche ein Spinarray der Länge $3m$ ist, wobei m die Anzahl der Klauseln in einem 3SAT-Problem darstellt, zurück auf eine Variablenbelegung der logischen Variablen des 3SAT-Problems abgebildet. Danach wird überprüft ob es nicht erfüllte Klauseln gibt. Ist dies der Fall, so wird für jede Variable einer nicht erfüllten Klausel überprüft, ob deren Negation bereits die Belegung *wahr* erhalten hat. Ist dies nicht der Fall, so kann der betrachteten Variable der Wahrheitswert *wahr* zugewiesen werden. In diesem Fall werden weitere Variablen dieser Klausel nicht mehr angesehen und es wird direkt mit der nächsten nicht erfüllten Klausel fortgefahren.

5 Evaluation

Wie in Abschnitt 2.1 dargelegt wurde, gibt es bei 3SAT-Instanzen signifikante Unterschiede bezüglich der Schwierigkeit der Lösungsfindung. Das Verhältnis aus Anzahl der Klauseln zur Größe des Variablenpools, im Folgenden K/V -*Verhältnis* genannt, bestimmt dabei die Schwierigkeit der Lösungsfindung. Unter Berücksichtigung dieses Sachverhalts werden im Folgenden zwei Versuchsreihen durchgeführt.

1 SAT w/ Phase Transition

0.5 SAT via WMIS on QA

1 Experimental Setup

5.1 2. Versuchsreihe

In der 2. Versuchsreihe werden mit Hilfe des ToughSAT-Generators jeweils 100 zufällig generierte 3SAT-Formeln für die K/V-Verhältnisse 0.2, 4.2 sowie 8.4 erzeugt. Alle 3SAT-Formeln bestehen aus 42 Klauseln. Damit stehen zur Erzeugung der Formeln beim K/V-Verhältnis von 0.2 insgesamt 210 Variablen, beim K/V-Verhältnis von 4.2 insgesamt 10 Variablen und beim K/V-Verhältnis von 8.4 insgesamt 5 Variablen zur Verfügung. Analog zur 1. Versuchsreihe werden hier alle 100 3SAT-Formeln der K/V-Verhältnisse 0.2 sowie 4.2 so gewählt, dass sie erfüllbar sind. Alle 3SAT-Formeln des K/V-Verhältnisses 8.4 sind so gewählt, dass sie nicht erfüllbar sind.

5.2 Versuchsdurchführung

Jedes 3SAT-Problem der 1. sowie der 2. Versuchsreihe wird zunächst auf ein WMIS-Problem reduziert (vgl. Kapitel ??). Die Knoten des durch diese Reduktion entstanden Graphen G_{SAT} werden mit dem Gewicht 1 belegt, sämtliche Kanten mit dem Gewicht 4. Die Gewichte sind hier beliebig gewählt, jedoch so, dass das Kantengewicht jeder Kante größer ist als das Minimum der Gewichte der Knoten die diese Kante verbindet (vgl. Behauptung in Kapitel 2.2). Anschließend werden aus diesem Graphen die linearen Gewichte der Ising-Energiefunktion berechnet (vgl. Gleichung ?? bzw. ??). Nun kann das Problem dem DWAVE-Quantenannealer übergeben werden, der zunächst ein Embedding dieses Problems auf dem Chimera-Graphen bestimmt und anschließend durch Quantum Annealing eine Lösung des Problems bestimmt. Für jede der in dieser Arbeit benutzten 3SAT-Formeln wird lediglich ein Embedding berechnet, welches dann für 1000 Annealingvorgänge benutzt wird. Man erhält also pro Formel insgesamt 1000 potentielle Lösungen.

1 Colourful Pictures (6.11, 6.12, 6.15, 6.19 + Numbers from Tab. 6.4)

1 Explaining the colourful pictures

1 Discussion & Conclusion (7.5.3)

6 Conclusion

References

1. Feld, S., Roch, C., Gabor, T., Seidel, C., Neukart, F., Galter, I., Maurer, W., Linnhoff-Popien, C.: A hybrid solution method for the capacitated vehicle routing problem using a quantum annealer. arXiv preprint arXiv:1811.07403 (2018)

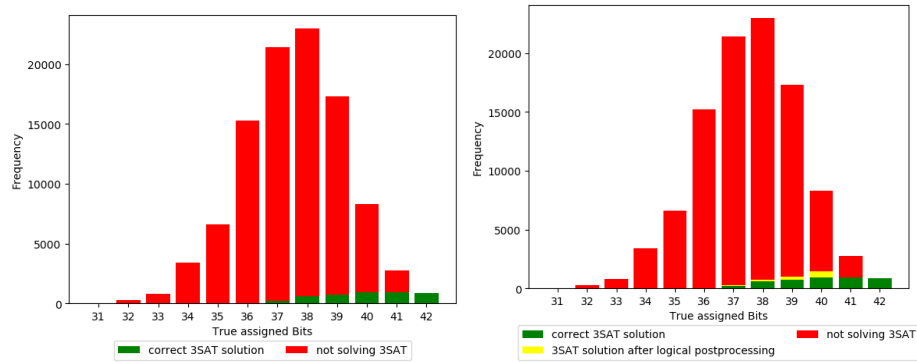


Fig. 2. Verteilung korrekter und nicht korrekter Antwortbitstrings ohne logisches Postprocessing (links) und mit logischem Postprocessing (rechts) für 3SAT-Instanzen bestehend aus 42 Klauseln und einem K/V- Verhältnis von 4,2.

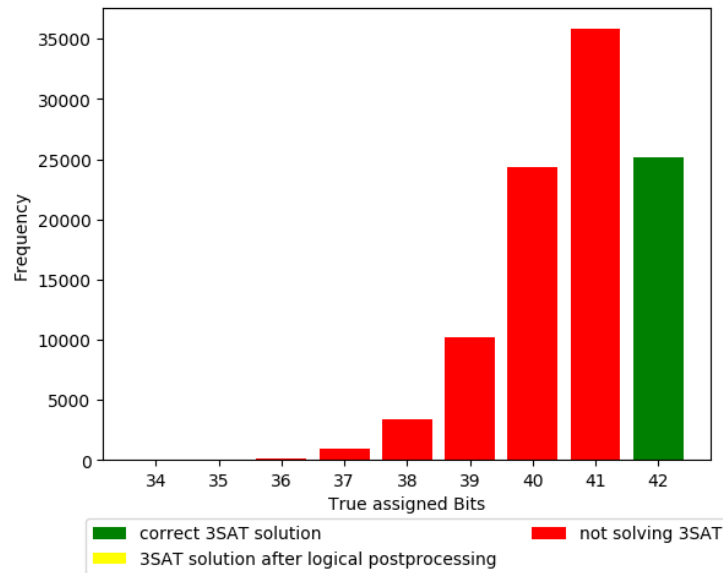


Fig. 3. Verteilung korrekter und nicht korrekter Antwortbitstrings mit logischem Postprocessing für 3SAT-Instanzen bestehend aus 42 Klauseln und einem K/V- Verhältnis von 4,2.

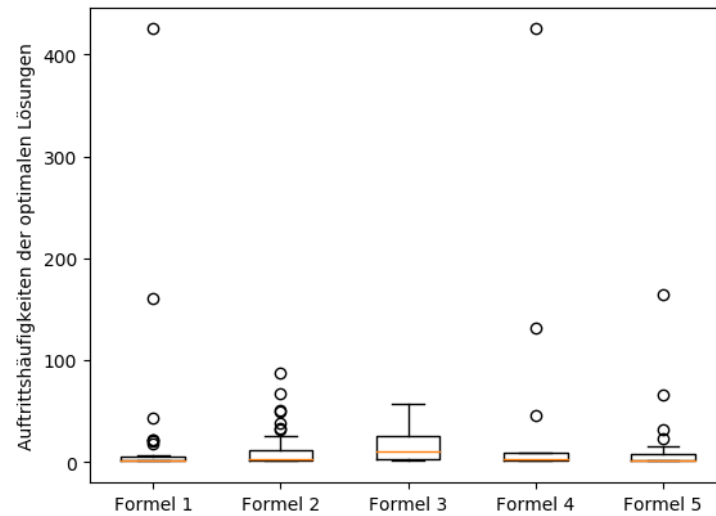


Fig. 4. Auftretshäufigkeiten verschiedener optimaler Lösungen für verschiedene 3SAT-Formeln.

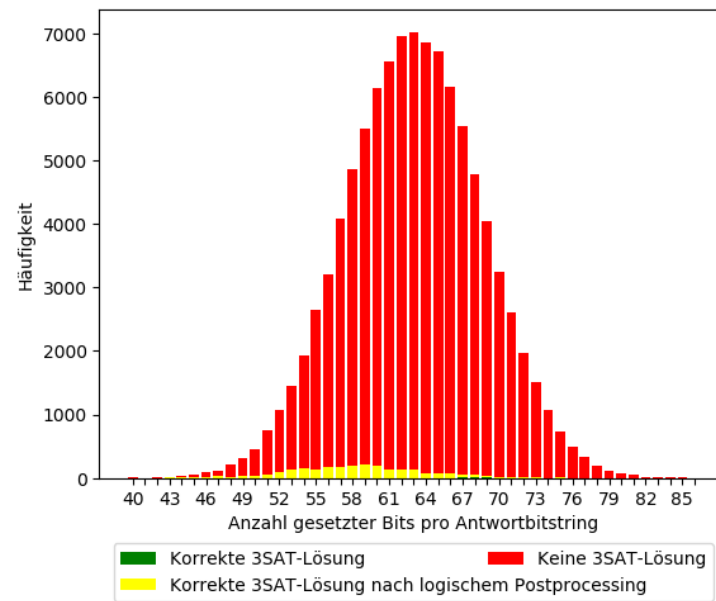


Fig. 5. Verteilung korrekter 3SAT-Lösungen und falschen Antworten unter Antwortbitstrings, die durch den naiven Zufallsgenerator erzeugt wurden, mit bestimmter Anzahl gesetzter Bits für 3SAT-Formeln bestehend aus 42 Klauseln.