Centre for
Data Analytics

Insight

Insight Centre for Data Analytics

From Backdoor Key to Backdoor Completability: Improving a Known Measure of Hardness for the Satisfiable CSP

Guillaume Escamocher, Mohamed Siala, Barry O'Sullivan

June 28, 2018

Context

Context

- Understanding instance hardness?

Context

- Understanding instance hardness?
- Not from a complexity point of view, but from solver point of view

Context

- Understanding instance hardness?
- Not from a complexity point of view, but from solver point of view
- The notion of backdoor

Backdoor

Backdoor

Intuition

- Hardness is concentrated in a small part of the instance.

Backdoor

Intuition

- Hardness is concentrated in a small part of the instance.
- The notion of backdoor isolates the area where hardness occurs in terms of variables.

Backdoor

Intuition

- Hardness is concentrated in a small part of the instance.
- The notion of backdoor isolates the area where hardness occurs in terms of variables.

Definition (Williams, Gomes, Selman 2003)

Let I be a satisfiable CSP instance, a backdoor is a subset B of variables such that there exists an assignment of the variables in B that makes it easy to find a solution for I.

Backdoor Key Fraction (BKF)

Backdoor Key Fraction (BKF)

- A **partial solution** is an instantiation that does not violate any constraint

Backdoor Key Fraction (BKF)

- A **partial solution** is an instantiation that does not violate any constraint

Definition (Ruan, Kautz, Horvitz 2004)

Let B be a backdoor for a satisfiable instance I.
Let $v \in B$ and let S a partial solution for $B \setminus \{v\}$.

- v is dependent: exactly one value a in the domain of v such that $S \cup \{a\}$ can be extended to a solution for I.
- key fraction of B: $\frac{\#\text{dependent variables in } B}{|B|}$

Backdoor Key Fraction (BKF)

- A **partial solution** is an instantiation that does not violate any constraint

Definition (Ruan, Kautz, Horvitz 2004)

Let B be a backdoor for a satisfiable instance I.
Let $v \in B$ and let S a partial solution for $B \setminus \{v\}$.

- v is dependent: exactly one value a in the domain of v such that $S \cup \{a\}$ can be extended to a solution for I.
- key fraction of B: $\frac{\#\text{dependent variables in } B}{|B|}$

Hypothesis

higher backdoor key fraction $\Leftrightarrow$ harder instance

Limitation of BKF

Limitation of BKF

There are two major cases where the backdoor key fraction is not useful.

Limitation of BKF

There are two major cases where the backdoor key fraction is not useful.

- Case when there is only one solution: $\implies$ BFK always 1.

Limitation of BKF

There are two major cases where the backdoor key fraction is not useful.

- Case when there is only one solution: $\Longrightarrow$ BFK always 1.
- Case when flipping the truth assignment of any variable in the backdoor and still extend the backdoor to a solution: $\Longrightarrow$ BFK always 0.

Characterizing Hardness in Backtracking Search

Characterizing Hardness in Backtracking Search

Search space The choices made by the solver.

Characterizing Hardness in Backtracking Search

Search space The choices made by the solver.

Solution space The choices made by the solver that lead to a solution.

Characterizing Hardness in Backtracking Search

Search space The choices made by the solver.

Solution space The choices made by the solver that lead to a solution.

Combining both

Characterizing Hardness in Backtracking Search

Search space The choices made by the solver.

Solution space The choices made by the solver that lead to a solution.

Combining both

- Large search space + small solution space
  $\implies$ A lot of choices made, few of them good
  $\implies$ The solver takes a long time to find a solution.

Characterizing Hardness in Backtracking Search

Search space The choices made by the solver.

Solution space The choices made by the solver that lead to a solution.

Combining both

- Large search space + small solution space
  $\Longrightarrow$ A lot of choices made, few of them good
  $\Longrightarrow$ The solver takes a long time to find a solution.
- Solution space close to the search space
  $\Longrightarrow$ Almost all the choices made are good
  $\Longrightarrow$ solver finds a solution quickly.

Completability Ratio

Completability Ratio

Definitions

Let B be a backdoor for I. Completability ratio of B is $\frac{\#completable}{\#partial}$, with:

- $\#partial$ the number of partial solutions for B.
- $\#completable$ the number of partial solutions for B that can be extended to a solution for I.

Backdoor Completability

Backdoor Completability

Definition

Backdoor completability of a satisfiable instance I:
average completability ratio of all minimal backdoors of I.

Backdoor Completability

Definition

Backdoor completability of a satisfiable instance I:
average completability ratio of all minimal backdoors of I.

Hypothesis

lower backdoor completability $\Leftrightarrow$ harder instance

Theoretical Justification

Theoretical Justification

What we want

A class $\mathcal{C}$ of instances, a solver A.

- $\mathcal{C}$ is tractable for $A \rightarrow$ backdoor completability of **all** instances in $\mathcal{C}$ is **high**.
- $\mathcal{C}$ is not tractable for $A \rightarrow$ backdoor completability of **some** instances in $\mathcal{C}$ is **low**.

Theoretical Justification

What we want

A class $\mathcal{C}$ of instances, a solver A.

- $\mathcal{C}$ is tractable for $A \rightarrow$ backdoor completability of **all** instances in $\mathcal{C}$ is **high**.
- $\mathcal{C}$ is not tractable for $A \rightarrow$ backdoor completability of **some** instances in $\mathcal{C}$ is **low**.

Scope of the Result

Scope of the Result

Instance classes

- $\mathcal{C}_1 \subset \mathcal{C}_2 \subset \ldots$ where a satisfiable instance is in $\mathcal{C}_p$ if the treewidth of its primal constraint graph is bounded by p.
- Treewidth is often used for tractability results.
- The union of all $\mathcal{C}_p$ is equal to the whole satisfiable CSP.

Scope of the Result

Instance classes

- $\mathcal{C}_1 \subset \mathcal{C}_2 \subset \ldots$ where a satisfiable instance is in $\mathcal{C}_p$ if the treewidth of its primal constraint graph is bounded by p.
- Treewidth is often used for tractability results.
- The union of all $\mathcal{C}_p$ is equal to the whole satisfiable CSP.

Solvers

- $A_1, A_2, \ldots$ where A_q is based on $(q, 1)$-consistency.
- The transition between tractability and hardness is sharp: for any class C_p and solver A_q:
  - $p \leq q \rightarrow A_q$ finds $\mathcal{C}_p$ trivial.
  - $p > q \rightarrow A_q$ fails to solve some instances in $\mathcal{C}_p$.

  so the distinction between low and high completability values should be clear.

Theorem

For any two integers p, q exactly one of the following is true:

1. $p \leq q$ and for every $I \in \mathcal{C}_p$, the backdoor completability of I is 1.
2. $p > q$ and there is an infinite number of instances in $\mathcal{C}_p$ with a backdoor completability exponentially low in the number of variables.

Theorem

For any two integers p, q exactly one of the following is true:

1. $p \leq q$ and for every $I \in \mathcal{C}_p$, the backdoor completability of I is 1.
2. $p > q$ and there is an infinite number of instances in $\mathcal{C}_p$ with a backdoor completability exponentially low in the number of variables.

Proof

1. From the construction of the solvers.
2. For each number n, we build an instance in $\mathcal{C}_p$ with more than n variables and a backdoor completability with regard to A_q asymptotically lower than $\frac{1}{2^n}$.

Experimental study

Experimental study

- Quasigroup with Holes
  - 1100 instances of order 22
  - Number of “holes” between 192 and 222: peak of difficulty at 204

Experimental study

- Quasigroup with Holes
  - 1100 instances of order 22
  - Number of “holes” between 192 and 222: peak of difficulty at 204
- Random CSP
  - 1200 instances with 60 variables and 1770 constraints
  - Average constraint tightness uniformly between 5% and 16%: peak of difficulty at 8%.

Experimental study

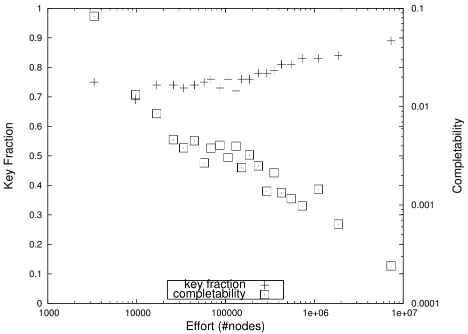
- Quasigroup with Holes
  - 1100 instances of order 22
  - Number of "holes" between 192 and 222: peak of difficulty at 204
- Random CSP
  - 1200 instances with 60 variables and 1770 constraints
  - Average constraint tightness uniformly between 5% and 16%: peak of difficulty at 8%.
- Mistral solver with default configuration

Experimental study

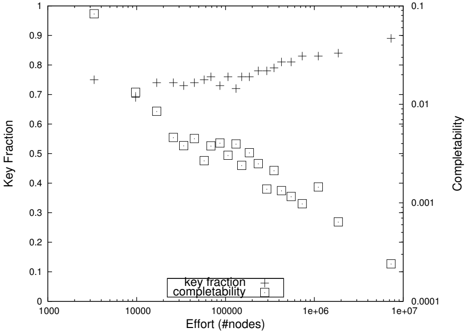
- Quasigroup with Holes
  - 1100 instances of order 22
  - Number of "holes" between 192 and 222: peak of difficulty at 204
- Random CSP
  - 1200 instances with 60 variables and 1770 constraints
  - Average constraint tightness uniformly between 5% and 16%: peak of difficulty at 8%.
- Mistral solver with default configuration
- Standard method to find backdoors

Quasigroup With Holes

Quasigroup With Holes



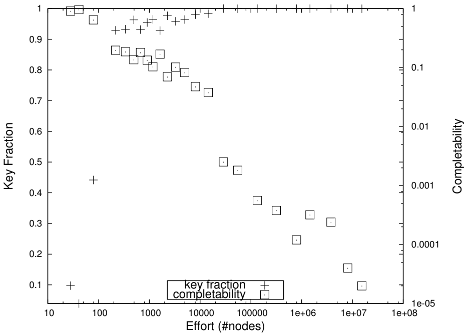
Quasigroup With Holes



		BD key fraction	BD completability
QWH	Pearson CC	.876	-.943
	RMSE	.053	.037
	MAE	.032	.029

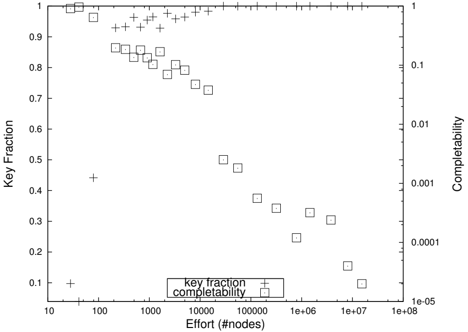
Pearson CC Pearson Correlation Coefficient
RMSE Root Mean Square Error
MAE Mean Absolute Error

Random CSP



Key Fraction
Completability
Effort (#nodes)
key fraction
completability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
1
0.1
0.01
0.001
0.0001
1e-05
10
100
1000
10000
100000
1e+06
1e+07
1e+08

Random CSP



		BD key fraction	BD completability
Random	Pearson CC	.590	-.975
	RMSE	.186	.051
	MAE	.158	.044

Pearson CC Pearson Correlation Coefficient

RMSE Root Mean Square Error

MAE Mean Absolute Error

Conclusions & Future Research

Conclusions & Future Research

Conclusion

- A better hardness measure compared to the backdoor key fraction
- Theoretical and experimental studies to support our claim

Conclusions & Future Research

Conclusion

- A better hardness measure compared to the backdoor key fraction
- Theoretical and experimental studies to support our claim

Future Research

- Predict hardness?

Conclusions & Future Research

Conclusion

- A better hardness measure compared to the backdoor key fraction
- Theoretical and experimental studies to support our claim

Future Research

- Predict hardness? Very expensive to compute

Conclusions & Future Research

Conclusion

- A better hardness measure compared to the backdoor key fraction
- Theoretical and experimental studies to support our claim

Future Research

- Predict hardness? Very expensive to compute : Could it be approximated?

Conclusions & Future Research

Conclusion

- A better hardness measure compared to the backdoor key fraction
- Theoretical and experimental studies to support our claim

Future Research

- Predict hardness? Very expensive to compute : Could it be approximated?
- Generate hard instances?

Conclusions & Future Research

Conclusion

- A better hardness measure compared to the backdoor key fraction
- Theoretical and experimental studies to support our claim

Future Research

- Predict hardness? Very expensive to compute : Could it be approximated?
- Generate hard instances?
- Design branching heuristics?

Thank you!