Review Of Unique Key Horn Functions

Kristóf Bérczi, Endre Boros, Ondřej Čepek, Petr Kučera, Kazuhisa Makino

Presented by: Timothy Deligero Presented Date: 4/7/20

Instructor: Charles Hughes

Introduction: What are Keys and Horns?

- What are *horn* functions?
  - Defined in the theory of special functions in mathematics as 34 distinct hypergeometric series of order two by Horn, hence its name, in 1931, and then was later corrected by Borngässer in 1933
  - Used and studied under a wide range of various topics over time due to its interesting structural and computational properties as a from of subclass of Boolean functions as a *pure Horn function*. [1]
  - In database theory, its *implicates*, $A \rightarrow v$, can serve as functional dependencies.

- What is a *key*?
  - In relational databases, is a set of attributes made to be assigned to a set of values.
  - Users can retrieve data based on certain conditions and requirements from a database's keys.
  - *Unique keys* can contain NULL values and be store multiple times across a database separately.

UCF

Proposal: Unique Key Horn Functions

- The authors of this research paper [1] presents the use of *unique key Horn functions* with the following concepts:
  - Focuses on Sperner hypergraphs B that form a set of minimal keys from a unique pure Horn function h_{B}. This type of hypergraph would be called a *unique key hypergraph*, with its corresponding Horn function called a *unique key Horn function*. [1]
  - Characterizations of a unique key hypergraphs and unique key Horn functions.
  - Unique key hypergraphs are proven to be co-NP-complete with edge sizes of 2 and other several cases of hypergraphs can be solved in polynomial time.
  - Connections between solutions to problems of minimal key generation from pure Horn functions and minimum target set selection of a graph, creating similar algorithms with polynomial delay.

Background and Related Work

- Horn functions have been used and studied under many various topics:
  - Directed hypergraphs in graph theory and combinatorics
  - Implication systems in machine learning
  - Database theory (e.g. relational databases as functional dependencies)
  - Lattices and closure systems in algebra and concept lattice analysis
  - Hydra functions using key Horn functions.

- Horn functions have very strong relation to database theory using their implications and CNF representations.

Unique Key Horn Functions: Terminology

- Definitions and notations established for *unique key graphs* and *unique key Horn functions*:
  - **Sperner hypergraphs (clutters)** – $\mathrm{G} = (V, E)$ or $\mathcal{B} \subseteq 2^V$
  - *Transversal* – $T \subseteq \mathrm{V}$ of $\mathcal{B}$ is when $T \cap B \neq \emptyset$, where $\forall B \in \mathcal{B}$.
  - *Independent set* – S of $\mathcal{B}$ is when $T = V \setminus S$ is a transversal of $\mathcal{B}$.
  - *Minimal transversals* – $\mathcal{B}^d$
  - *Family of independent sets* – $\mathcal{B}^*$
  - *Subhypergraph* – $\mathcal{B}_S = \{B \in \mathcal{B} \mid B \subseteq S\}$, where $S \subseteq V$.
  - *Projection of a hypergraph* – $\mathcal{B}^S = min'l\,\{S \cap B \mid B \in \mathcal{B}\}$, where $S \subseteq V$.
  - Union of hyperedge of a hypergraph – $\cup\, \mathcal{B}$ (i.e. $\cup\, \mathcal{B} = \bigcup_{B \in \mathcal{B}} B$).
  - *Forward chaining closure* – $F_h(S) = \{u \in V \mid S \rightarrow v$ is an implicate of $h\}$.
  - *Set of minimal keys* - $\mathcal{K}(h)$ for a Horn function CNF h.

Unique Key Horn Functions: Characterizations

- **Lemma 1.** *For a Sperner hypergraph* $\mathrm{B} \subseteq 2^V$ *and subset* $S \subseteq V$ *we have* $(\mathrm{B}_S)^d = \left(\mathrm{B}^d\right)^S$ *and* $(\mathrm{B}^S)^d = \left(\mathrm{B}^d\right)_S$. [1]

Well known established logic for hypergraphs.

- **Lemma 2.** *Let* $\mathrm{B} \subseteq 2^V$ *be a Sperner hypergraph and* $h : \{0,1\}^V \to \{0,1\}$ *be a pure Horn function such that* $h \leq \Phi_{\mathrm{B}}$. *Then* $K(h) \neq \mathrm{B}$ *if and only if there exists an implicate* $A \to v$ *of* h *and a minimal transversal* $T \in \mathrm{B}^d$ *such that* $T \cap A = \emptyset$ *and* $v \in T$. [1]

Proof. Apply properties established by the definitions and notations of unique key graphs and unique key Horn functions.

- **Lemma 3**. *Let* $\mathrm{B} \subseteq 2^V$ *be a Sperner hypergraph and* $h : \{0,1\}^V \to \{0,1\}$ *be a pure Horn function such that* $h \leq \Phi_{\mathrm{B}}$. *Then* $K(h) \neq \mathrm{B}$ *if and only all implicates* $A \to v$ *of* h *with* $A \in \mathrm{B}^*$ *we have* $v \in (V \backslash \mathrm{A})(\cup \mathrm{B}^{V \backslash \mathrm{A}})$. [1]

- **Lemma 4.** *Let* $\mathrm{B} \subseteq 2^V$ *be a Sperner hypergraph and define* $\Psi = \{A \to v \mid A \in \mathrm{B}^*, v \notin \cup \mathrm{B}^{V \backslash \mathrm{A}}\}$. *Let* φ *be a set of clauses of the form* $A \to v$ *that are not implicates of* Φ_{B}. *Then* $K(\varphi \wedge \Phi_{\mathrm{B}}) = \mathrm{B}$ *if and only if* $\varphi \subseteq \Psi$. [1]

Proof. Apply proofs of Lemmas 1 and 2.

UCF

Unique Key Horn Functions: Characterizations

- **Theorem 5.** *For a Sperner hypergraph* $\mathrm{B} \subseteq 2^V$*, the pure Horn function* $h = \Phi_{\mathrm{B}}$ *is the only one with* $K(h) = \mathrm{B}$ *if and only if for all* $T \in \mathrm{B}^d$ and $v \notin T$ *there exists* $T' \in \mathrm{B}^d$ *such that* $T' \neq T$ *and* $T' \subseteq T \cup \{v\}$. [1]

Proof. Apply Lemma 4 and use arbitrary values for if direction $A \in \mathrm{B}^*$ *and* $v \in \cup\, \mathrm{B}^{V \setminus \mathrm{A}}$, and for the only if direction $T \in \mathrm{B}^{\mathrm{d}}$ and $v \notin T$, where $A = V(T \cup \{v\})$.

- **Corollary 6.** *The cuts of a loopless matroid form a unique key hypergraph*. [1]

Proof. The set of minimal traversals B^{d} can serve as cut-sets and base-sets of matroids, and a loopless matroid would compliment Theorem 5.

- **Remark 7.** *The conditions of* **Theorem 5** *can be checked in polynomial time if* B^d *can be generated in (input) polynomial time from* B*. For example, if* B *is 2-monotone or forms the set of bases of a matroid.*

Unique Key Graphs: Complexity

- **Theorem 8.** *A graph* $G = (V, E)$ *is unique key if and only if for every maximal independent set* $I \subseteq V$ *and vertex* $v \in I$ *there exists a vertex* $u \notin I$ *that is an individual neighbor of* v. [1]

Proof. The set of minimal keys are complement to the family of independent sets, especially maximal independents sets, resulting in minimal traversal representing minimum vertex covers.

- **Theorem 9.** *A CNF* Φ *is not satisfiable if and only if the graph* G_Φ *is unique key.* [1]

Proof. Use maximal independent sets and its relation to neighboring vertices within a graph to justify the cases where a CNF Φ is satisfiable and not a unique key graph, creating a co-NP-complete solution.

- **Corollary 10.** *Deciding if a hypergraph is unique key is co*-NP-*complete already for hypergraphs of dimension 2.* [1]

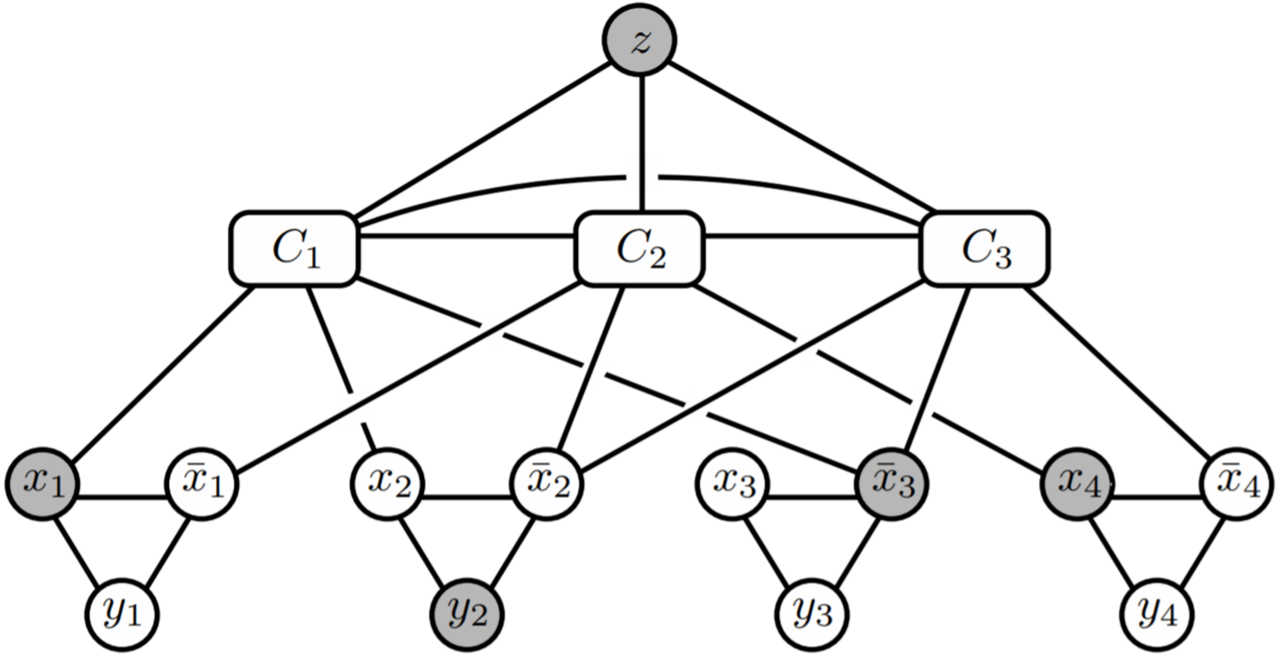


Fig. 1. The graph G_Φ corresponding to CNF formula $\Phi = (x_1 \vee x_2 \vee \bar{x}_3) \wedge (\bar{x}_1 \vee \bar{x}_2 \vee x_4) \wedge (\bar{x}_2 \vee \bar{x}_3 \vee \bar{x}_4)$. Grey vertices form a maximal independent set corresponding to a satisfying truth assignment. Note that z has no individual neighbor. [1]

Unique Key Graphs: Examples

- **Theorem 11.** *A bipartite graph $G = (V, E)$ without isolated vertices is unique key if and only if E is a perfect matching*. [1]

Proof. Every maximal independent set of a unique key graph contains exactly one end vertex for every edge in a unique key graph, creating a perfect matching bipartite graph.

- **Theorem 12.** *For graphs of bounded treewidth, it is possible to decide in linear time if a graph is a unique key graph*. [1]

- **Corollary 13.** *For graphs of bounded clique-width, it is possible to decide in linear time if a graph is a unique key graph.* [1]

Proof. Given that there are maximal independent sets in a unique key graph, it can be solved using predicates to satisfy the conditions for a unique graph and the treewidth value can also represent the clique-width of a CNF.

UCF

Unique Key Graphs: Examples

- **Theorem 14.** *Let* $G = (V, E)$ *be a graph, and assume that the size of the largest induced matching of* G *is bounded by a constant. Then there is an efficient algorithm to decide if* G *is a unique key graph*. [1]

Proof. The induce matching is proportional to the size of the graph which also affects the number of independent sets. If the size is bounded by a constant, then checking all independent sets for a unique key graph takes polynomial time.

MIN-KEY and MIN-TSS Problems

- **Lemma 15.** D_Φ *is strongly connected*. [1]

Proof. Show that there exists a path from a key K_3 to minimal keys K_1 and K_2, while showing that the distance between K_3 and K_2 is smaller than the distance between K_2 and K_1.

- **Theorem 16.** *Given a pure Horn CNF* Φ*, we can generate all minimal keys of* Φ *with polynomial delay.* [1]

Algorithm:

1. Given that D_Φ is strongly connected, then all out-neighbors will be generated from the minimal keys that are already generated, starting from a minimal key which is generated by greedily leaving out elements from V. [1]
2. Store the minimal keys in a LIFO queue.
3. Generate the out-neighboring vertices/values of the top element of the queue and add the new ones to the queue.
4. Output the top element of the queue.
5. Repeat Steps 3) and 4) until all minimal keys are generated.

MIN-KEY and MIN-TSS Problems

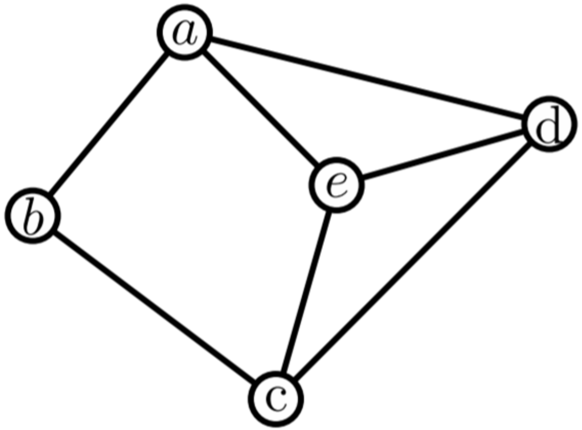
- **Theorem 17.** *The* MIN-TSS *problem with constant thresholds is polynomial-time reducible to the* MIN-KEY *problem*. [1]

Proof. Assume a given undirected graph G with threshold values for activation. The generation algorithm from Theorem 15 matches the activation process for a minimum target set selection. This results in that the key $K \subseteq V$ is a target set of G if and only if it is a unique key graph. Process is shown in Fig. 2.

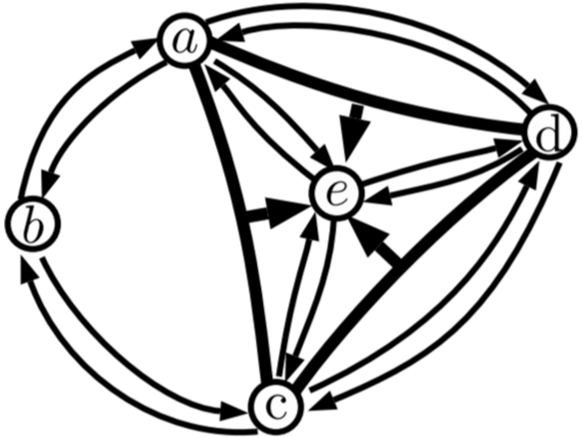
- **Theorem 18.** *The* MIN-KEY *problem with constant thresholds is polynomial-time reducible to the* MIN-TSS *problem.* [1]

Proof. Construct an undirected graph $G = (V', E)$ based on a given pure Horn CNF and its variables while also adding gadgets for each clause, so that minimal keys sets can also be minimal targets sets with no changes in size for the minimal keys. Process is shown in Fig. 3

- **Corollary 19.** *Given a graph $G = (V, E)$ and constant thresholds $t : V \to \mathbb{Z}_+$, we can generate all minimal target sets of G with polynomial delay.* [1]

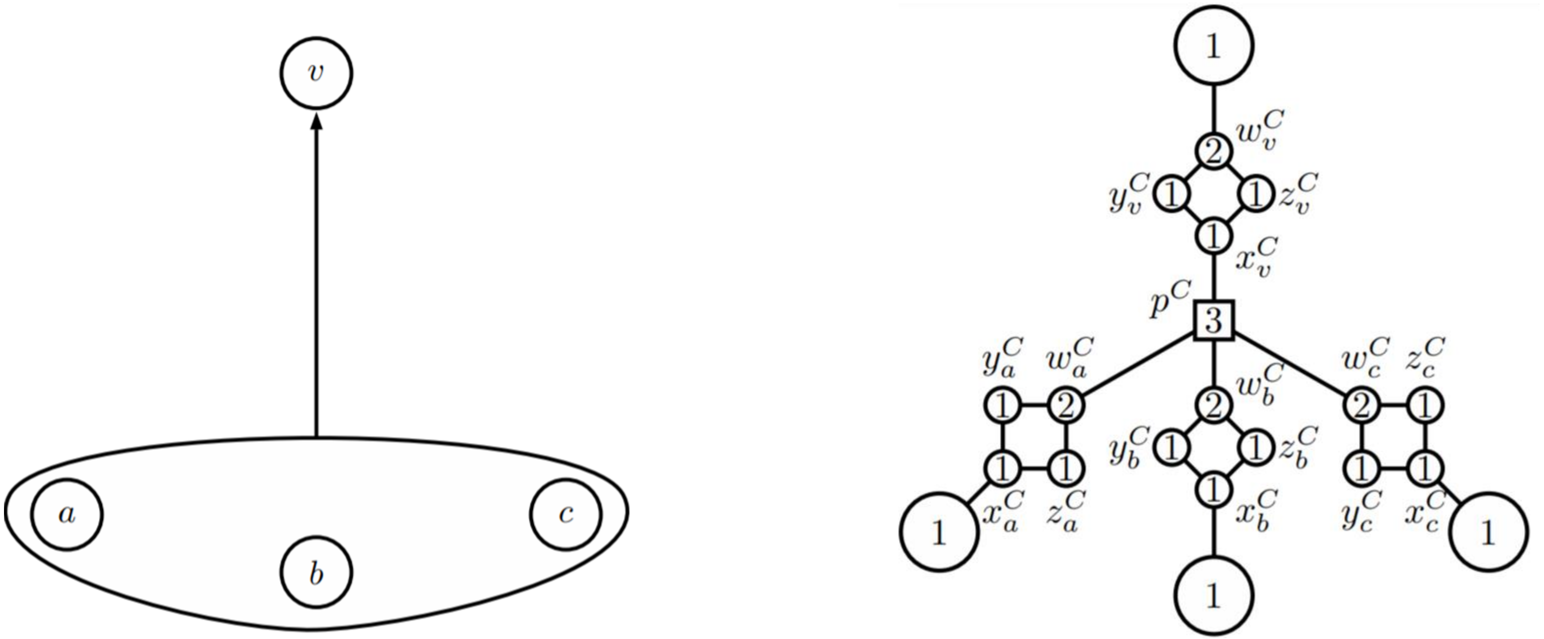


(a) Instance of MIN-TSS problem. The thresholds are $t(a) = t(b) = t(c) = t(d) = 1$ and $t(e) = 2$.



(b) Construction of Ψ_G. Thick hyperedges represent clauses containing three variables.

Fig. 2. An illustration of Theorem 17. The CNF associated to G is $\Psi_G = (b \to a) \wedge (e \to a) \wedge (d \to a) \wedge (a \to b) \wedge (c \to b) \wedge (b \to c) \wedge (d \to c) \wedge (e \to c) \wedge (a \to d) \wedge (c \to d) \wedge (e \to d) \wedge (\{a, c\} \to e) \wedge (\{a, d\} \to e) \wedge (\{c, d\} \to e)$. [1]



(a) A pure Horn clause $C = A \rightarrow v$, where $A = \{a, b, c\}$.

(b) The gadget and threshold values corresponding to $C = A \rightarrow v$.

Fig. 3. An illustration of Theorem 18. Note that the size of the graph G is polynomial in the length of the input. [1]

Discussion and Summary

- **Major Contributions**:
  - Effective elaboration of definitions and notations used for Lemmas, Theorems, and Corollaries.
  - Presented cases of representations of unique key graphs present new different models to hypergraphs.
  - Algorithms are both effective and efficient for both MIN-KEY and MIN-TSS problems and show valid similarities.
  - The new reduction using gadgets presented for the MIN-TSS can potentially be used for various problems to reduce to the MIN-TSS problem to prove they can be solved with polynomial delay.

- **Weaknesses and Criticisms**:
  - Little variety in the models that can be represented as hypergraphs, or specifically unique key graphs.
  - Only presents one example to associate with the MIN-KEY problem and does go into detail of the polynomial delay dependent on the inputted CNF.

- **Future research and extensions**:
  - Full survey paper on the versatility of pure Horn CNFs.
  - Extended research on different types of graphs that can be presented as hypergraphs or unique key graphs under certain conditions, as well as apply the MIN-KEY problem to other various problems outside of MIN-TSS.
  - Look into unbounded MIN-TSS problem and observe if it can be solve in polynomial delay when associated with the MIN-Key problem.
  - Provide software implementation on the presented algorithms for effectiveness and efficiency analysis.

UCF

Conclusion

- The authors of the research paper [1] define unique key graphs and their corresponding unique key functions with multiple Lemmas, Theorems, and Corollaries.
- Provide proof of Sperner hypergraphs with edge sizes of 2 co-NP-complete as well as several cases of graphs that can be proven to be unique key graphs in polynomial time.
- Present algorithms for the MIN-KEY and MIN-TSS problems, as well as show strong relations between the two problems by showing they are polynomial time-reducible to each other.
- Overall research contributes to graph theory and database theory, as well as pure Horn functions in general, leaving potential for future extensions to the presented topics.

References

[1] K. Bérczi, E. Boros, O. Čepek, P. Kučera, and K. Makino, “Unique key Horn functions,” Feb. 2020.

Questions and Answers

Questions?