*URL http www elsevier nl locate entcs volume html pages*

*Symbolic State Exploration*

*Fabio Somenzi*

*Department of Electrical and Computer Engineering University of Colorado*

*Boulder CO USA*

*Abstract*

*The exploration of the state space of the model is at the heart of model checking Symbolic algorithms often use Binary Decision Diagrams for the representation of sets of states and have to pay attention to the size of the BDDs We review the techniques that have been proposed for this task in the framework of closed sequences of monotonic functions*

*Introduction*

*Formal veri cation techniques like model checking explore the state space* of the system to be veri ed To combat the exponential growth of the number of states with the number of state variables symbolic algorithms for state ex ploration refrain from dealing explicitly with individual states Instead they manipulate the characteristic functions of sets of states In the late eight ies the foundations for symbolic state exploration were laid by the work of McMillan and co workers Coudert Berthet and Madre and Pixley

*The common trait of those early works was the formulation of decision*

*procedures for classes of properties in terms of state reachability and the solu tion of the reachability problem by breadth rst search of the state transition* graph with graph and state sets represented by Binary Decision Diagrams

*BDDs*

*In cases of properties like EF p there exists an execution of the system* during which p holds at some time the decision procedure had to establish the reachability of a state labeled p from other states This could be done by moving either forward or backward in the state transition graph In the case of properties like EG p there exists an execution of the system during which p holds at all times the decision procedure had to establish the existence of a cycle labeled by p reachable from some other states along a path also labeled

*This work was supported in part by SRC contract DJ*

*c Published by Elsevier Science B V Open access under* [*CC BY-NC-ND license*](http://creativecommons.org/licenses/by-nc-nd/3.0/)*.*

*by p The rst problem translates into the computation of a least xpoint while the second entails the computation of a greatest xpoint*

*In general the logics commonly used in model checking can be expressed* in terms of calculus and in practice a model checker spends most of its time computing xpoints State exploration and computation of xpoints are therefore closely related in the context of symbolic model checking Fixpoint computation is founded on Tarksi s theorem which states that the least and greatest xpoints of monotonic functions over nite lattices can be computed by repeated application of the functions starting from either the least elements of the lattices for least xpoints or the greatest elements of the lattices for greatest xpoints The powerset of the set of states is the lattice on which the formally monotonic functions expressing the xpoints of interest in model checking operate The repeated application of functions maps quite naturally onto symbolic breadth rst search

*Symbolic breadth rst search however is not always the most e cient ap proach to state exploration Consider the computation of the states reachable* from a designated state s At each iteration a lower bound to the xpoint is computed Speci cally the initial lower bound is fsg and at the k th it eration of the symbolic depth rst search algorithm all states such that the shortest path from s to them has length k are added to the previous bound to create a new one The sequence of lower bounds the iterates of the xpoint computation obviously form a monotonic sequence However the sizes of the BDDs that represent the iterates seldom form a monotonic sequence In many cases the BDD sizes reach a maximum that may be several times the value at convergence An example is shown in Figure The reason for such a behavior of the BDD sizes is that the states at a given maximum distance from the initial states may not be well clustered in the the boolean space This leads to ine cient representation in terms of BDDs As the computation pro ceeds the gaps between states may be lled by other states leading to better BDDs Sometimes large intermediate sizes are the result of sub optimal BDD variable orders and can be remedied by applying dynamic reordering Other times there is no good order and the growth in BDD size triggers several expensive reorderings to no avail

*The problem with intermediate results being much larger than the nal* results occurs also within each step of the breadth rst search procedure Each step computes either the successors or the predecessors of a set of states by a series of conjunctions and quanti cations The BDDs for the intermediate results are often much larger than those encountered at the end This problem compounds the one of the sizes of the xpoint iterates and often prevents or seriously hinders completion of a model checking run

*Symbolic breadth rst has other problems besides the size of the interme diate results On the one hand even an optimal search strategy may produce* excessively large BDDs either at some intermediate point or at the end Computing an approximation to a xpoint may be the only alternative to

5e+06

4.5e+06

4e+06

number of BDD nodes in Reached

3.5e+06

3e+06

2.5e+06

2e+06

1.5e+06

1e+06

500000

0

BFS

0 5 10 15 20 25 30 35 40

number of iterations

*Fig Size of the BDDs for xpoint iterates*

*giving up on the exact computation On the other hand obtaining the entire*

*xpoint may be a waste of time as in the case in which it is su cient to es tablish reachability of one state from some other states Once a path is found that connects the desired states the computation can be terminated The ideal search strategy proceeds towards the target states possibly disregarding states that are closer to the origin*

*In summary the size of intermediate results the impossibility to compute the xpoints of some large problems exactly and the desire to focus the search in the directions that cause early termination has motivated the development of alternatives to breadth rst search for xpoint computation It goes with out saying that the opposite of breadth rst search depth rst search is usually much worse in the context of symbolic state exploration Therefore the methods that have been devised combine elements of depth rst search in a basic scheme that is still quite close to breadth rst in that it tries to manipulate large sets of states at once*

*This paper reviews some of the search strategies that have been proposed in a uniform framework This framework closed sequences of monotonic functions is presented in Section It is a natural extension of the results on xpoints of monotonic functions to the case in which multiple functions need to be evaluated Section formulates several search strategies in terms of closed sequences and Section concludes*

*Closed Sequences of Monotonic Functions*

*De nition Let L be a nite lattice and let T f kg be a nite* set of monotonic functions L L For l l L let l l be the least upper bound of l and l For a nite sequence over T let be the function L L obtained by composing all the functions in in the order speci ed by the sequence Let j j be the length of sequence Let designate the concatenation of sequences and We say that is closed if for i k we have i

*If is closed is a xpoint of every i T If is closed over T then* it is closed over any subset of T

*Lemma Let be a nite sequence over T Then is monotonic*

*Proof By induction on j j If j j there is nothing to prove Suppose* j j i Let i be the pre x of length i of Then by the inductive hypothesis and the monotonicity of i

*l li li l i i l i i l Therefore l l*

*Lemma Let be a nite sequence over T Then*

*Proof From and the monotonicity of*

*Lemma Let be a nite closed sequence over T Let be any nite sequence over T Then*

*Proof By Lemma and closure of*

*Corollary Let be two closed sequences over T Then*

*Proof By Lemma we have*

*and*

*Lemma Let p be a sequence of monotonic functions over L Let p be a sequence of functions over L such that i i for i p Then*

*Proof By induction on p For p we have Let i i be the pre x of length p of Assumeii Then*

*i i i i i i*

*where the rst inequality comes from i i and the second comes from* the inductive hypothesis and the monotonicity of i

*Lemma Let be a nite closed sequence over T Let Then*

*Proof From we get by monotonicity and closure of*

*Theorem Let be a closed sequence over T Let P T be the pointwise least upper bound of all the functions in T Then*

*X X*

*Proof By de nition of and closure of we have*

*X X*

*T T*

*Hence is a xpoint of which implies X X Let be the* sequence obtained by repeating j j times Then by Lemma we have

*X X*

*Hence is the least xpoint of*

*Theorem only requires monotonicity of the functions in T In reacha bility analysis we can assume that x x but this additional assumption is* unnecessary Notice that the function over the boolean algebra B f a b g de ned by x a shows that for a monotonic function it is possible to have

*x x and also x not comparable to x*

*However there may not exist closed sequences for a given T Consider* T f g with x and x Let be a sequence over T If ends with then Conversely if ends with then

*De nition A function L L is upward if x x for all x L*

*De nition An in nite sequence n over T is fair if for every n and for j k there exists i n such that i j*

*Theorem If all functions in T are upward then every fair sequence over T has a nite pre x that is closed*

*Proof Let i be the pre x of length i of a fair sequence Since i is upwardii From the niteness of L it follows that there exists an n such that for i nin Let mj be the smallest i n*

*such that i j j k Then mj mj m nn*

*j*

*Hence n is closed*

*Lemma Let L L be a monotonic function and let x x Then x x*

*Proof By Lemma x x We prove by induction that x*

*x For the basis we observe that x For the inductive step we* assume i x We then have

*i i x x x x*

*Theorem Let P T be the pointwise least upper bound of all the*

*functions in T Let T f kg with i x x i Let be a fair*

*sequence over T Then has a nite pre x that is closed and such that*

*x*

*Proof By construction i is upward Hence by Theorem has a pre x*

*closed over T By Theorem and Lemma x x*

*Theorem Let be a closed sequence over T Let L L be de ned by x x for Let i be the sequence obtained from by inserting after i Then i is closed and*

*i*

*Proof Let We have*

*Hence is closed over T fg By Lemmai Let i*

*i Let be the sequence obtained from by replacing i with i From Lemma it follows that Fromi we then geti and nallyi and that is closed over T*

*i*

*Except for Lemma which is its own dual the remaining results have* duals that apply to greatest xpoint computations Here we state explicitly the main results

*De nition For l l L let l l be the greatest lower bound of l and l We say that is closed if for i k we have i A function L L is downward if x x for all x L*

*Theorem Let be a closed sequence over T Let Q T be the pointwise greatest lower bound of all the functions in T Then*

*X X*

*Theorem If all functions in T are downward then every fair sequence over T has a nite pre x that is closed*

*Theorem Let Q T be the pointwise greatest lower bound of all*

*the functions in T Let T f kg with i x x i Let be a fair sequence over T Then has a nite pre x that is closed and such that*

*x*

*Theorem Let be a closed sequence over T Let L L be de ned by x x for Let i be the sequence obtained from*

*by inserting after i Then i is closed and*

*i*

# *Search Strategies*

*In symbolic state exploration we assume we are given a monotonic function that operates on sets of states of the model to be veri ed For instance the* computation of EF p entails the computation of Z p Pre Z where Pre Z computes all the predecessors of the states in Z In this case Z p Pre Z This section shows how for di erent search strategies a set T f kg is derived from and how closed sequences over T are constructed

*Disjunctive Partitioning*

*To alleviate the problem of large intermediate results during one iteration* of the xpoint computation Cabodi et al have proposed the disjunctive partitioning of In the simplest form of partitioning the set T f kg is obtained by identifying a set of state variables such that cofactoring or restricting the transition relation according to all their possible assignments leads to i s with small BDDs The functions are upward and are applied in round robin fashion This gives fair sequences for which Theorems and guarantee convergence to the desired xpoint Cabodi et al allow the partitioning to change at each image computation Their approach has been extended by Narayan et al to allow among other things more exible sequencing of the i s while still guaranteeing fairness

*High Density Reachability Analysis*

*The states that are expanded in one iteration of xpoint computation form the* frontier set Breadth rst search always tries to expand all unexpanded states

*and possibly more High Density HD Reachability Analysis on the* other hand may use a subset of the unexpanded states as frontier set BDD approximation functions are used to that e ect Because of subsetting the function applied at the i th iteration i satis es the condition i Furthermore all i are upward However there is no guarantee that

*Pi i Hence HD Reachability Analysis resorts to dead end computations*

*to guarantee convergence Let j be the function applied when the j th dead*

*end computations is performed j is upward and guarantees that j Z Z*

*only if j Z Z Hence if a sequence over T f ig f jg is closed then*

*P T By Theorems and convergence to Z is guaranteed*

*An approach related to HD Reachability Analysis is Saturated Simulation*

*in which however there is no provision to guarantee convergence*

*Symbolic Guided Search*

*Symbolic guided search applies hints to the transition relation of the sys tem to be veri ed that are meant to enable only some modes of operation or restrict addresses to some values or sequences The hints are conjoined* to the transition relation to yield the set T f kg All the functions in T are upward Each function is applied until it yields no further states Repeated application of the same function enables the use of the frontier set To guarantee convergence always includes as last hint the hint that gives

*k This choice tends to keep the number of iterations needed to reach*

*convergence low When switching from j to j the guided search algo rithm cannot use the frontier set This my lead to an expensive computation especially when switching to k For this the dead end computation al gorithm of HD Reachability Analysis is used An alternative approach which* may be useful to avoid some of the most expensive dead end computations is to iterate the application of f k g This alternative approach requires

*P i k i Symbolic guided search has been applied to model checking*

*of some LTL properties and also to CTL*

*Approximate Reachability Analysis*

*Approximate Reachability Analysis can be used to compute a superset of* reachable states of a system This nds applications in sequential optimiza tion as well as in model checking The computation of an upper bound to the reachable states is accomplished by decomposing the system in sub machines The decomposition produces a set T f kg such that

*S T One can then use Corollary to prove that any fair sequence*

*over T produces the same approximation of the reachable states As long*

*as the approximation algorithm is otherwise the same for all sequences The* interesting fact is that some sequences are much more e ciently computed than others

*Conclusions*

*Closed sequences of monotonic functions provide a natural framework to ex press search strategies used in symbolic state exploration These strategies* address the limitations of pure breadth rst search that cause the BDDs pro duced by the model checking algorithms to grow too large In this paper we have shown in particular how High Density Reachability Analysis Disjunctive

*Partitioning Symbolic Guided Search and Approximate Reachability Analy sis can be formulated in a uniform way as computations of closed sequences*

# *Acknowledgements*

*I thank Roderick Bloem Jim Kukula In Ho Moon Kavita Ravi and Tom Shiple for providing inspiration for this work and comments on early versions of the manuscript*

# *References*

*R Bloem K Ravi and F Somenzi E cient decision procedures for model checking of linear time logic properties In N Halbwachs and D Peled editors Eleventh Conference on Computer Aided Veri cation CAV pages Springer Verlag Berlin LNCS*

*R E Bryant Graph based algorithms for boolean function manipulation IEEE Transactions on Computers C August*

*J R Burch E M Clarke K L McMillan D L Dill and L J Hwang Symbolic model checking states and beyond In Proceedings of the Fifth Annual Symposium on Logic in Computer Science June*

*G Cabodi P Camurati L Lavagno and S Quer Disjunctive partitionining and partial iterative squaring An e ective approach for symbolic traversal of large circuits In Proceedings of the Design Automation Conference pages*

*Anaheim CA June*

*G Cabodi P Camurati and S Quer Improved reachability analysis of large nite state machines In Proceedings of the International Conference on Computer Aided Design pages Santa Clara CA November*

*H Cho G D Hachtel E Macii B Plessier and F Somenzi Algorithms for approximate FSM traversal based on state space decomposition IEEE Transactions on Computer Aided Design December*

*H Cho G D Hachtel E Macii M Poncino and F Somenzi Automatic state space decomposition for approximate FSM traversal based on circuit analysis IEEE Transactions on Computer Aided Design December*

*O Coudert C Berthet and J C Madre Veri cation of sequential machines based on symbolic execution In J Sifakis editor Automatic Veri cation Methods for Finite State Systems pages Springer Verlag LNCS*

*O Coudert C Berthet and J C Madre Veri cation of sequential machines using boolean functional vectors In L Claesen editor Proceedings IFIP International Workshop on Applied Formal Methods for Correct VLSI Design pages Leuven Belgium November*

*M Dam CTL and ECTL as fragments of the modal calculus Theoretical Computer Science*

*E A Emerson and C L Lei E cient model checking in fragments of the propositional mu calculus In Proceedings of the First Annual Symposium of Logic in Computer Science pages June*

*M Fujita Y Matsunaga and T Kakuda On variable ordering of binary decision diagrams for the application of multi level logic synthesis In Proceedings of the European Conference on Design Automation pages Amsterdam February*

*K L McMillan Symbolic Model Checking Kluwer Academic Publishers Boston MA*

*I H Moon J Y Jang G D Hachtel F Somenzi C Pixley and J Yuan Approximate reachability don t cares for CTL model checking In Proceedings of the International Conference on Computer Aided Design pages San Jose CA November*

*I H Moon J Kukula T Shiple and F Somenzi Least xpoint approximations for reachability analysis In Proceedings of the International Conference on Computer Aided Design San Jose CA November To appear*

*A Narayan A J Isles J Jain R K Brayton and A L Sangiovanni Vincentelli Reachability analysis using partitioned ROBDDs In Proceedings of the International Conference on Computer Aided Design pages November*

*C Pixley A computational theory and implementation of sequential hardware equivalence In E M Clarke and R P Kurshan editors Computer Aided Veri cation pages American Mathematical Society Association for Computing Machinery*

*K Ravi K L McMillan T R Shiple and F Somenzi Approximation and decomposition of decision diagrams In Proceedings of the Design Automation Conference pages San Francisco CA June*

*K Ravi and F Somenzi High density reachability analysis In Proceedings of the International Conference on Computer Aided Design pages San Jose CA November*

*K Ravi and F Somenzi E cient xpoint computation for invariant checking In Proceedings of the International Conference on Computer Design Austin TX October To appear*

*K Ravi and F Somenzi Hints to accelerate symbolic traversal In CHARME September To appear*

*A Tarski A lattice theoretic xpoint theorem and its applications Paci c Journal of Mathematics*

*J Yuan J Shen J Abraham and A Aziz On combining formal and informal veri cation In O Grumberg editor Ninth Conference on Computer Aided Veri cation CAV pages Springer Verlag Berlin LNCS*