ITP 20002 Discrete Mathematics, Fall 2020

Homework 2. Set Representations and SAT-based Puzzle Solving (Revised at Oct 06)

Shin Hong

1. Overview

This homework asks you to implements two versions of integer set data structures, one with array (Section 2.1) and the other with bitvector (Section 2.2), and construct an automatic solver of the Minesweeper puzzle (Section 3) using the set data structure and SAT solver Z3[1].

Set data structures are designed to contain a set of integer values. Array-based data structure represents a membership of an integer by containing it in an integer array, thus, set operations perform searches, addition, and removal of integer values on the integer array. Bitvector-based data structure marks the membership of its elements on a Boolean vector (calling it bitvector). Set operations for bitvector-based data structure involve bitwise operations on the internal bitvector.

Minesweeper is a grid-based puzzle where a player is asked to find out all cells with mines based on the given clues. You must design a scheme to transform a given puzzle instance into a corresponding propositional logic formula and construct a puzzle solving program in C using your own set data structure implementation and SAT solver Z3.

The submission deadline is **11:00 PM, 14 October** (Wed). Your submission must include the implementations of the two versions of set data structures, a Minesweeper puzzle solving program and a report on all tasks of Homework 2.

2. Set data structure

You can find the baseline code from the `homework2` branch at `https://github.com/hongshin/DiscreteMath`. Directory `array` contains the code for array-based data structure, and `bitvect` for bitvector-based data structure.

2.1. Array-based data structure

An integer set can be represented with an array where an integer is contained if and only if the integer is the member of the set. It is possible to determine whether or not an integer is contained in a set by checking if the integer exists in the array. Note that, although members are arranged in a certain order, the ordering does not represent any information of the set.

You can find the structural definition of an array-based set representation `intset` in `array/intset.h`. The `n_elems` field is to represent the current number of members. The `elems` field is to point an integer array whose length is `n_size`. The interface of the set operations is declared also at `array/intset.h`. The definitions of the operations to construct and deconstruct `intset` objects are given at `array/intset.c`.

2.2. Bitvector-based data structure

For a finite universe with total n distinct values, a set can be represented as a vector of n bits. Each bit is mapped uniquely with a specific value, such that the truth value of a bit indicates whether a corresponding value is contained in a set.

The structure of the bitvector-based set representation is defined as `intset` in `bitvect/intset.h`. The `univ` field is to point an integer array that contains all values of integers, and the n_univ field indicates the size of the univ integer. A bitvector-based set object must be constructed with the given values of `univ` and `n_univ`. You can assume that no duplication exists in `univ`, and the values in `univ` never chance once it is used for constructing set objects. The `bitvect` field points to an array to represent the bitvector where the n-th bit is to mark whether or not `univ[n]` is contained in the set. The `n_elems` field represents the number of members contained in the set.

The interface of `intset` is declared at `bitvect/intset.h`. Note that the types of some operations are difference from the array-based version. The definitions of object construction and deconstruction operations are given at `bitvect/intset.c`.

2.3. Set operations

Both versions of the set data structure must provide the following operations:

- $size(S)$: returns the number of distinct elements contained in the given set S (i.e., cardinality).
- $add(S, v)$: update the given set S by adding the given integer v (if v was not in S).
- $remove(S, v)$: update the given set S by removing the given integer v (if v was in S)
- $contains(S, v)$: determine whether the given integer v is contained in the given set S, or not.
- $equals(S_1, S_2)$: determine whether the given two sets S_1 and S_2 are equivalent.
- $union(S_1, S_2)$: create a new set that is the union of the given two sets S_1 and S_2.
- $intersection(S_1, S_2)$: create a new set that is the union of the given two sets S_1 and S_2.
- $difference(S_1, S_2)$: create a new set that is the union of the given two sets S_1 and S_2.
- $subsets(S, k)$: create an array of all k-size subsets of the given set S.
- $powerset(S)$: create an array of all subsets (i.e., power set) of the given set S.

The details of operation requirements are given as comment at `array/set.c` and `bitvet/set.c`. These operations must be implemented to satisfy all given requirements, and as adhere to the existing definitions of the structure and the operations. Also, you can find some samples of operation use cases at the test case code (`test.c`).

[1] The latest version of Z3 can be found at https://github.com/Z3Prover/z3

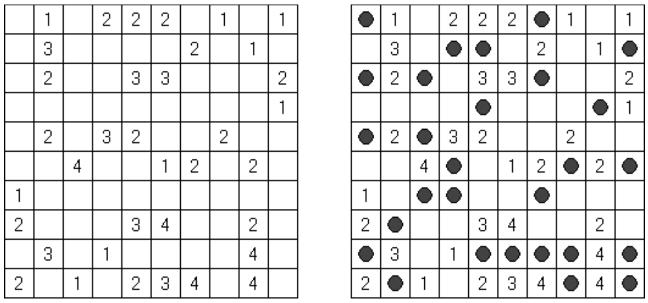

Figure 1. Minesweeper Puzzle Example

3. Minesweeper puzzle

A Minesweeper puzzle[2] consists of a $n \times n$ square grid where one or multiple mines are installed invisibly at some of its cells. A player of this puzzle is asked to find out all cells with mines based on the given numbers labeled on some cells. A number given at a cell indicates the number of the surrounding cells with mines (thus, a number is between 0 and 8). Also, a number implies that the cell where the number is labeled does not have any mine. Figure 1 shows an example of a Minesweeper game. The left figure depicts the initial setting of the game before the play, and the right one shows all locations of the mine installed cells with black circle.

As similar for Homework 1, you are asked to model a puzzle instance as a propositional logic formula such that an assignment that satisfies the formula represents a solution of the puzzle instance (the formula should be unsatisfiable if a puzzle instance has no solution). Based on this modelling, you must construct a program that uses a SAT solver Z3 to find a solution of a puzzle instance given as input. The description is not repeated, but you are asked to take the same as given in Section 3 of Homework 1.

4. Requirements

4.1. Set data structure implementation

You are asked to complete the missing parts of the two set data structure implementations. You are restricted to add code to where the `/*TODO*/` comment exists. Note that You are not allowed to change for all the other parts.

For the bitvector-based set data structure, you need to use proper bitwise operators[3] to implement set operations.

You can build each version by `make` (see `Makefile` in each directory). Your program must be compatible with GCC 5.4.0 or a higher version. As your program will be tested on the peace server[4], it is recommended to check if your program is working well on it.

You can run tests on your program by `make test`. This command prints "`Pass`" when your implementation passes all test cases given in `test.c`, or you will see that an assertion violation happens when your program unexpectedly behaves. Note that your program will be tested with more test cases at evaluation.

4.2. Minesweeper puzzle

An input is given from the standard input. The first line contains one positive integer n, the length of one side of the puzzle grid. n does not exceed 15. The second to the $(n+1)$-th lines define the initial grid with the number label. Each of these lines has n words separated by one or multiple white spaces. The j-th word at the $(i+1)$-th line describes the cell of the i-th row and the j-th column. A word is either a number between 0 and 8, or '?'; a number indicates that the corresponding cell is labeled with that number and a question mark indicates that the corresponding cell has no label. The following is the input text for Figure 1:

```
10
? 1 ? 2 2 2 ? 1 ? 1
? 3 ? ? ? ? 2 ? 1 ?
? 2 ? ? 3 3 ? ? ? 2
? ? ? ? ? ? ? ? ? 1
? 2 ? 3 2 ? ? 2 ? ?
? ? 4 ? ? 1 2 ? 2 ?
1 ? ? ? ? ? ? ? ? ?
2 ? ? ? 3 4 ? ? 2 ?
? 3 ? 1 ? ? ? ? 4 ?
2 ? 1 ? 2 3 4 ? 4 ?
```

You can assume that a well-formed input is given to your program.

Your program must print out the result to the standard output. If a solution is found, all locations of mines must be displayed in a human readable form. Your program must notify that there is no solution if it is impossible to satisfy all constraints at the same time.

Your puzzle-solving program must use one of your set data structure implementations. Explain in your report for what purpose a set data structure is used for constructing a puzzle solver.

4.3. Report

You must write one report that covers all aspects of Homework 2 (including two set data structures, and a puzzler solver implementations). You must use a given report template and your report must not exceed 2 pages. All contents must be written in English (none of your Korean sentences will be counted in evaluation). Your report must be converted to PDF for submission.

You must detail in your report the following aspects of your solution:

(1) your idea and details on accomplishing each design and implement task
(2) structure of implementations
(3) validation results
(4) open discussion, for instances, lessons learned, observations, further investigation, suggestions of possible extension or improvements, challenges, limitations of your submission, etc.

In addition, as a part of discussion, discuss in which cases an array-based or a bitvector-based data structure has more benefits over the other version.

Evaluation will be primary based on your report (not your implementation), thus, you must try best to deliver all your results with best presentations in the reports.

5. Submission

Submit an archive of all your results (e.g., tar or zip file) by 11:00 PM, 14 October (Wed) via Hisnet. Your submission must include (1) reports in PDF, (2) source code files of two versions of set data structure, and (3) a C source code file of a Minesweeper puzzle solver. The submission deadline is strict, and no late submission will be accepted.

[2] This example is taken from http://cross-plus-a.com

[3] c.f. https://en.wikipedia.org/wiki/Bitwise_operations_in_C

[4] peace.handong.edu