Homework 3: due by 11:59pm on Thursday, October 4, 2012
(Total: 100 points)
Instructor: Vivek Sarkar
Co-Instructor: Ran Libeskind-Hadas

All homeworks should be submitted using the submission system at http://cs.hmc.edu/submit. In case of problems using the system, please email your homework to cs181ehelp@cs.hmc.edu.

Honor Code Policy: All submitted homeworks are expected to be the result of your individual effort. You are free to discuss course material and approaches to problems with your other classmates, the teaching assistants and the professor, but you should never misrepresent someone else’s work as your own. If you use any material from external sources, you must provide proper attribution.

1 Written Assignment (50 points total)

Please submit your solution to this assignment in a plain text file named hw3_written.txt in the submission system.

1.1 Phasers and Atomic Integers (25 points)

Consider the incomplete HJ program below in Listing 1 followed by a desired inter-task synchronization pattern in Figure 1.

1. (15 points) Complete the missing phaser declarations and phaser registrations in lines 3, 8, 11, 14, to obtain a complete HJ program that implements the inter-task synchronization pattern in Figure 1.

2. (10 points) Assume that a is a reference to a Java AtomicInteger object initialized to zero, and that each of steps A() L() include the following code, “int n = a.getAndAdd(1);”. What are the possible values that variable n can receive in step H()?

```
finish {
    // INSERT MISSING PHASER DECLARATIONS AND INITIALIZATIONS BELOW

    // INSERT MISSING PHASER REGISTRATIONS BELOW
    async phased ( --------------------------- ) { //Task T1
        A(); next; B(); next; C(); next; D();
    } // Task T1
    async phased ( --------------------------- ) { //Task T2
        E(); next; F(); next; G(); next; H();
    } // Task T2
    async phased ( --------------------------- ) { //Task T3
        I(); next; J(); next; K(); next; L();
    } // Task T3
} // finish
```

Listing 1: Incomplete HJ program with missing phaser declarations and registrations
1.2 HJ isolated statements vs. Java atomic variables (25 points)

Many applications use Pseudo-Random Number Generators (PRNGs) as in class IsolatedPRNG in Listing 2. The idea is that the seed field takes a linear sequence of values obtained by successive calls to the nextInt() method as shown in line 7. The use of the HJ isolated statement in lines 5–8 ensures that there will be no data race on the seed field if nextSeed() is called in parallel by multiple tasks. The serialization of the isolated statement instances will determine which task obtains which seed in the sequence.

```java
class IsolatedPRNG {
  private int seed;
  public int nextSeed() {
    int retVal;
    isolated {
      retVal = seed;
      seed = nextInt(retVal);
    }
    return retVal;
  } // nextSeed()
} // IsolatedPRNG
```

Listing 2: Concurrent PRNG implemented using isolated statement

Since the isolated statement is not available in Java, the version below in class AtomicPRNG (Listing 3) attempts to implement the same functionality by using Java atomic variables instead.

1. (15 points) Assuming a scenario where nextSeed() is called by multiple tasks in parallel on the same PRNG object, state if the implementation of AtomicPRNG.nextSeed() has the same semantics as that of IsolatedPRNG.nextSeed(). If so, why? If not, why not?

2. (10 points) Why is the “while (true)” loop needed in line 5 of AtomicPRNG.nextSeed()? What would happen if the while loop was removed?
class AtomicPRNG {
    private AtomicInteger seed;
    public int nextSeed() {
        int retVal;
        while (true) {
            int retVal = seed.get();
            int nextSeedVal = nextInt(retVal);
            if (seed.compareAndSet(retVal, nextSeedVal)) return retVal;
        } // while
    } // nextSeed()
    . . . // Definition of nextInt(), constructors, etc.
} // AtomicPRNG

Listing 3: Concurrent PRNG implemented using Java’s AtomicInteger class

2 Programming Assignment (50 points total)

2.1 Habanero-Java Setup

See Section 2.1 of the Homework 0 handout for instructions on setting up a Habanero-Java installation for
use in this homework. You are encouraged to read Section 8 of the handout for Module 1 to understand the
different schedulers available in the HJ system.

The input files needed for this assignment can be found in [hw_3.zip]. See Section 3 of this handout for
available command-line options when compiling and executing HJ programs. In general, the command-line
interface (rather than DrHJ) is recommended for performance measurements, since DrHJ incurs additional
overhead on the machine by running in a separate JVM from the one used to execute your application.

2.2 Constraint Satisfaction Search algorithms

Constraint-satisfaction problems arise frequently in several applications areas including puzzle-solving and
engineering design. These problems are computationally intensive and well suited for speedup through
parallel processing. This assignment explores parallelization of constraint-satisfaction search algorithms that
use forward checking. It is well known that some form of look-ahead, as in forward checking, reduces the
sequential execution time for the application thereby making it a more credible candidate for parallelization
than simple backtracking. This assignment will focus on the use of constraint-satisfaction search in puzzle-
solving, with n-queens and Sudoku puzzles as two use cases.

An intermediate state of a constraint-satisfaction search is characterized by a partial Problem State in which
some variables have a single assigned value, and a Feasible Value Table (FVT), that provides a set of possible
values for the remaining free variables. If the set becomes empty for any variable, then it implies that no
feasible solution can be derived from the given intermediate state. If an FVT has exactly one value per
variable, then it can be combined with the partial Problem State to obtain a complete Problem State.

In the sequential code given to you, you can find the constraint-satisfaction search code in method search() of
ConstraintSatisfaction.hj, which is also shown in Listing 4 below. In this method, parameter state
(of type ProblemState) contains the partial problem statement on entry, parameter curVar identifies the
current variable to be explored in the search, and parameter fvt contains the FVT on entry. Line 2 checks
if we are examining the last variable, in which case no further forward checking is needed; instead, state
and fvt together identify one or more solutions that can be added to the set of feasible solutions. The loop
body in lines 7–11 of Listing 4 is then repeated for each feasible value v of curVar in fvt. This can be seen
in line 9 which constructs a new state in which curVar = v. The call to forwardCheck() in line 10 prunes
fvt to obtained a reduced newFvt that removes values of variables that are not feasible in conjunction with
curVar = v. If newFvt is null it means that no feasible solution is possible for curVar = v. Otherwise, the
public void search(ProblemState state, int curVar, FVT fvt) {
    if (curVar == fvt.getNumVars())
        problem.addSolution(state); // feasible solution found
    else {
        Iterator<Integer> itr = fvt.getValues(curVar).iterator();
        while (itr.hasNext()) {
            Integer v = itr.next();
            ProblemState newState = state.copy();
            newState.setValue(curVar, v);
            FVT newFvt = forwardCheck(curVar, v, fvt);
            if (newFvt != null) search(newState, curVar+1, newFvt);
        } // while
    } // if
} // search()

Listing 4: search() method in ConstraintSatisfaction.hj

public FVT forwardCheck(int curVar, Integer curVal, FVT fvt) {
    FVT newFvt = new FVT(fvt.getNumVars());
    for (int freeVar = curVar+1; freeVar < fvt.getNumVars(); freeVar++) {
        Iterator<Integer> itr = fvt.getValues(freeVar).iterator();
        while (itr.hasNext()) {
            Integer v = itr.next();
            if (problem.isConsistent(curVar, curVal, freeVar, v))
                newFvt.addValue(freeVar, v);
        } // while
        if (newFvt.getValues(freeVar).size() == 0) return null;
    } // for
    return newFvt;
} // forwardCheck()

Listing 5: forwardCheck() method in ConstraintSatisfaction.hj

recursive call to search() in line 11 explores feasible values for later variables.

Listing 3 contains the sequential code for method forwardCheck() of ConstraintSatisfaction.hj, which was called in line 10 of Listing 4. Line 3 iterates through the remaining free variables starting with curVar+1, and line 5 iterates through the values v that can be taken by freeVar. The essence of forwardCheck() is captured by the call to problem.isConsistent(curVar, curVal, freeVar, v) in line 7, which checks if the assignment of v to freeVar is consistent with the assignment of curVal to curVar. If so, the assignment of v to freeVar is added to newFvt (otherwise, it is not added). Line 10 checks if the set of feasible values for freeVar is empty. If so, a null value is returned instead of newFvt since no feasible solution exists at this point.

2.3 The Sequential Constraint-Satisfaction Solver

We have provided you an implementation of a sequential constraint-satisfaction search algorithm in the hw_3.zip file containing the following:

1. IConstraintSystem.hj — this interface defines the methods that a game/puzzle should implement in order for it to be solvable by our solver.

2. ProblemState.hj — this class represents the state of the game.
3. ConstraintSatisfaction.hj — this file contains two public methods, `solve()` and `parallelSolve()`. `solve()` contains a sequential implementation of the constraint-satisfaction search algorithm. `parallelSolve()` calls `parallelSearch()` which is currently empty, and will need to be filled in by you in parts 1 and 2 of your assignment (Section 2.4). The main programs also perform a comparison test on the results returned by `solve()` and `parallelSolve()`, which should pass after you’ve completed the assignments (but which fails for `parallelSolve()` in the version provided).

4. NQueens.hj — client solver for the NQueens problem.

5. Sudoku.hj — client solver for Sudoku, with a Reader to read input problems from the disk.

6. NQueensMain.hj — main program that starts and validates the results of the NQueens solver.

7. SudokuMain.hj — main program that starts and validates the results of the Sudoku solver.

Many puzzles can be represented by a set of rules that, applied on the current state of the puzzle, decide what are the possible actions that can be performed, which lead to a new puzzle state (with an assignment of values to a subset of free variables), thereby making them amenable to constraint-satisfaction search. This homework focuses on NQueens and Sudoku as two examples of such puzzles.

Alternative approaches to solving the NQueens puzzle have already been studied in class. The default size used by NQueensMain for this problem is \( n = 12 \).

Sudoku is a popular puzzle game that requires players to fill in missing numbers from 0 to N-1 on a square \( N \times N \) board, taking into account the following constraints:

- No square contains more than a number
- Every number appears only once on each column of the board.
- Every number appears only once on each row of the board.
- Every number appears only once in each individual region of the board. Regions are usually rectangular areas of size \( \sqrt{N} \times \sqrt{N} \) size.

Although Sudoku games are usually 9×9 with 3×3 regions, as in the 9x9.txt file, there are also variations that take larger board sizes as input, such as 16x16.txt with 4x4 regions. If 9x9 boards use the digits 1..9 to fill the board, larger sizes use 1..9, A, B, C, etc for the same purpose. Furthermore, some variations of Sudoku allow for multiple solutions, and the solver provided indeed finds all the possible solutions. If a cost function is specified, the solver must find the “cheapest” solution. The supplied serial solver already does that for you; however, it is your responsibility to keep the solver working in all case. The default input used by SudokuMain.hj is 9x9-2-multisol.txt.

2.4 Your Assignment: Parallel Constraint-Satisfaction Search

Your assignment is to design and implement a parallel algorithm for constraint-satisfaction search, using the provided sequential implementation as a starting point. Your homework deliverables are as follows.

1. **Computation of all solutions (15 points)**
   Create a new parallel version of ConstraintSatisfaction.hj that is designed to achieve the smallest execution time on the parallel machine (with 2 or more cores) that you use for this homework. You will be graded on the correctness and speedup of your parallel version. You can focus your attention on parallelizing the `search()` method. Keep in mind that the call to `problem.addSolution()` in line 3 of Listing 4 is not thread-safe i.e., it can lead to interference and data races if it is called multiple times (with different solutions) in parallel.
Your solution should work for any constraint-satisfaction problem, but you should test it with \texttt{NQueensMain.hj}, both for correctness and for achieving the best performance that you can relative to the sequential version. The \texttt{addSolution()} method in \texttt{NQueensCS.hj} accumulates all solutions into a single array. (It does not attempt to find the best solution according to some cost function.)

Please place all files related to this solution in a zip file named “\texttt{hw3\_part1.zip}”. (You are allowed to add or modify methods in any class, so long as your implementation still works for any constraint-satisfaction problem, and the \texttt{parallelSearch()} method returns the same results as \texttt{search()}.)

2. \textbf{[Computation of lowest-cost solution (20 points)]}

Create a new parallel version of \texttt{ConstraintSatisfaction.hj} that is designed to achieve the smallest execution time to return a single solution with lowest cost (so long as at least one feasible solution exists). While one approach is to simply reuse the solution from Part 1 above and return the solution with lowest cost, you can be smarter and reduce the work done by pruning the exploration of partial solutions that are guaranteed to never lead to a solution lower than the current best solution. You will be graded on the real speedup achieved by both your sequential and parallel versions relative to this brute-force approach. (The speedup will come from a combination of algorithmic improvements and from parallel computing.)

Your solution should work for any constraint-satisfaction problem with a cost function, but you should test it with \texttt{SudokuMain.hj}, both for correctness and for achieving the best performance that you can. The \texttt{addSolution()} method in \texttt{SudokuCS.hj} stores the best solution found (if any) using a (artificial) cost function that corresponds to the single \texttt{BigInteger} value obtained by scanning the digits from left-to-right and top-down in the solution. accumulates all solutions into a single array. As written, \texttt{getCost()} only provides a meaningful cost if the solution is a final Sudoku solution. You are welcome to extend to provide useful bounds for partial solutions, if you so choose.

Please place all files related to this solution in a zip file named “\texttt{hw3\_part2.zip}”. (You are allowed to add or modify method definitions in any class, so long as your implementation still works for any constraint-satisfaction problem with a cost function, and the \texttt{parallelSearch()} method returns the same result as \texttt{search()}.)

3. \textbf{[Homework report (15 points)]}

You should submit a brief report summarizing the design of your parallel algorithms in Parts 1 and 2 above, explaining why you believe that each implementation is correct and data-race-free.

Your report should also include the following measurements for both parts 1 and 2:

(a) Performance of the sequential version with the default input

(b) Performance of the parallel version with the default input, executed with the “\texttt{-places 1:1}” option to run with one worker and the “\texttt{-places 1:k}” option to run with \(k\) workers (where \(k\) is the number of cores available in the machine that you used.)

Please submit your report in a plain text file named \texttt{hw3\_report.txt} in the submission system. If you prefer to submit a PDF file, that’s okay too. (Send email to cs181ehelp@cs.hmc.edu with your report, if you have problems submitting a PDF file.)

3 Compile-time and Runtime options for Habanero Java

These options are described assuming that HJ programs are compiled and executed using the command-line interface. See DrJava’s Preferences panes to select some of these options when using DrJava.

To compile an HJ program, \texttt{Foo.hj}, type “\texttt{hjc [options] Foo.hj}”. The following command-line options are currently available for \texttt{hjc} (type “\texttt{hjc -help}” for a summary):
-racedet Enable race-detection when program executes (off by default, only works for basic async and finish constructs)

-rt s Compile program for work-sharing runtime (on by default, supports all HJ constructs)

-rt h Compile program for work-stealing runtime with help-first policy (off by default, only works for basic async and finish constructs)

-rt w Compile program for work-stealing runtime with work-first policy (off by default, only works for basic async and finish constructs)

-classpath ⟨path⟩ Search path for class files

-sourcepath ⟨path⟩ Search path for hj source files (must include -classpath if this option is used)

-destdir ⟨path⟩ Set the location where output classes from hjc should be placed

-version Print version number of the HJ compiler. Please include this version number when reporting any problems to comp322-staff.

To execute a compiled HJ program, Foo.hj, type “hj [options] Foo [args]”, where args are the command-line arguments for your program. The following command-line options are currently available for hj (type “hj -help” for a summary):

-places ⟨p⟩:⟨w⟩ Set number of places and workers per places. The default value is p = 1 and w = number of processors in the system.

   For now, you will be working with 1 place, so you can use -places 1:n to run an HJ program with n workers (for n cores).

-perf=true Output abstract execution metrics when program executes (off by default, only works for basic async and finish constructs)

-fj Use Fork-Join variant of work-sharing runtime (off by default, assumes that program has been compiled for work-sharing execution)

-version Print version number of the HJ runtime system. Please include this version number when reporting any problems to comp322-staff.

-mx ⟨size⟩ set max heap size, e.g., -mx 8000M sets the max heap size to 8GB

-classpath ⟨path⟩ Search path for class files

-J⟨arg⟩ Pass ⟨arg⟩ directly to Java runtime (use with caution)