Homework 2

Assigned: Wednesday, February 6, 2008

Due: Monday, February 18, 2008

Total value: 100 points + 5 points bonus.

Notes: This homework must be done individually. *If you receive help from anyone, you must clearly acknowledge it.* Always acknowledge sources of information (URL, book, class notes, etc.). Please inform instructor quickly about typos or other errors.

Content

---

1 Terminology

(Total 8 points)

Define the following terms, cite your references:

1. Complete graph
2. Clique
3. Direct acyclic graph

4. Efficient algorithm
5. Mathematical programming
6. Numeric constraint
7. Tractable problem
8. Triangulated graph

2 Configuration and Design: Lot development

(Total 12 points)
From Nadel (1989) via Dechter

The map in Figure 1 shows eight lots available for development. Five developments are to be

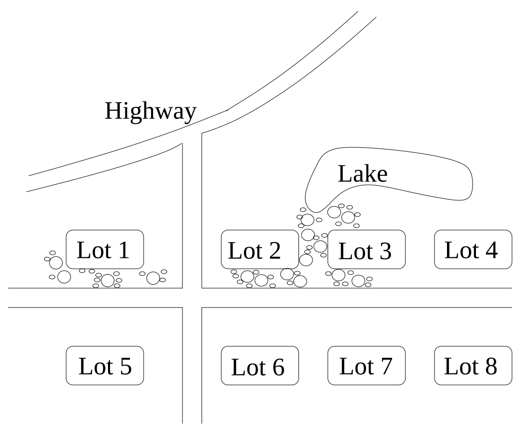

Figure 1: *Development map.*

located on these lots: a recreation area, an apartment complex, a cluster of 50 single-family houses, a large cemetery, and a dump site. Assume the following information and conditions:

- The recreation area must be near the lake.
- Steep slopes must be avoided for all but the recreation area.
- Poor soil must be avoided for developments that involve construction, namely, the apartments and the houses.
- Because it is noisy, the highway must not be near the apartments, the houses or the recreation area.
- The dump site must not be visible from the apartments, the houses, or the lake.

- Lots 3 and 4 have poor soil.
- Lots 3, 4, 7, and 8 are on steep slopes.
- Lots 2, 3, and 4 are near the lake.
- Lots 1 and 2, are near the highway.

Formulate this problem as a CSP. Define the variables, their domains, the constraints and the query. Clearly state any assumptions you make that are not listed above, otherwise

3 Reduction of 3SAT into a CSP

(Total 30 points)

1. Formulate 3SAT as a CSP. (15 points)
   Indications: Your formulation should be as general as possible and should represent each of the elements of the 3SAT and its question in the terminology of a CSP. Consider X, the set of Boolean variables of a 3SAT instance. What are the values that a variable can take? Use this to define the variables of the CSP and their values. A clause is a disjunction of literals. How to represent a clause in the CSP formalism? A 3SAT sentence is a conjunction of clauses. How is the sentence represented in the CSP formalism? Finally, state how the question of 3SAT is reduced as a question to the CSP and prove that a solution to 3SAT exists if and only if a solution to the corresponding CSP exists.
2. What is the arity of the constraints of the resulting CSP? (3 points)
3. As a direct application of your reduction, transform the following 3SAT problem into a CSP. Specify the variables, their domains, define the constraint in extension, and draw the corresponding constraint graph: (6 points)
   $$(c_1 \vee c_2 \vee c_3) \wedge (c_2 \vee c_3 \vee c_4) \wedge (\neg c_1 \vee c_5) \wedge (c_1 \vee c_4 \vee c_5)$$
4. Knowing how a 3SAT clause (which is a disjunction of at most 3 literals) is represented in the CSP, how do you propose to represent a clause of SAT (which has an arbitrary number of literals in the clause)? (6 points)

4 N-Queen Problem as a CSP

(Total 30 points)

Part 1: Consider the 4-queens problem where each queen is associated with a row and can be assigned to any column in the row.

1. Define this problem as a CSP. Specify the variables and their domain, and each binary constraint by 'extension.' (2 points)

2. Define a binary constraint C_{Q_i,Q_j} between two variables Q_i and Q_j by 'intension.' (4 points)

3. What is the size of this CSP? (1 point)

4. Draw the constraint graph. (3 points)

5. Arc-consistency of a binary constraint C_{Q_i,Q_j} between two variables Q_i and Q_j ensures that every value for the variable Q_i has a support (at least one consistent value) in the domain of Q_j and vice-versa. Run manually arc-consistency on the 4-Queens problem. Can you remove any value? At the end of the operation the CSP is said to be arc-consistent. (4 points)

6. Arc-consistency is also called 2-consistency because it considers all combinations of two variables at the same time. Let's consider all combinations of 3 variables at the same time and let's check whether or not every value in the domain of a given variable has a support in the domain of the two other variables (simultaneously). If it does not, the value can be removed. Can you remove any value? (Alternatively stated, a CSP is said to be strong 3-consistent, if for every combination of consistent values for two variables, one can find a value in the domain of any third variable such that the constraints between the three variables are satisfied.) (4 points)

7. What can you conclude about the effectiveness of consistency algorithms for the 4-Queens problem? (2 points)

Part 2: Alternative modeling of the N-queens problem.

Adapted from Bernard Nadel'1990.

1. Since there can be no more than one queen per right (downward-sloping) diagonal we could associate a variable with each diagonal. According to this model, how many variables are necessary to model the N-queens problem as a CSP? Give your answer in terms of N. (3 points)

2. Each variable taking its value to indicate in which square of the corresponding diagonal the queen is, there are less queens than diagonals and some diagonals must contain no queen. Find a way to define the domains of these variables to allow the modeling of the statement: "a diagonal may contain no queen." (4 points)

3. What is the size of the CSP resulting from this model (state in terms of $N = 4$)? (3 points)

4. **Bonus:** Express the size of the CSP as a general expression in N. (5 points)

5 Latin square

(Total 10 points)

Adapted from of Daphne Koller, Stanford University.

A Latin Square is a $N \times N$ array filled with colors, in which no color appears more than once in any row or column. Finding a solution to a 4×4 Latin Square can be formulated as a CSP, with a variable for each cell in the array, each having a domain of $\{r, g, b, y\}$, and a set of constraints asserting that any pair of cells appearing in the same row must have different colors, and that any pair of cells appearing in the same column must have different colors.

1 **r**	2 **g**	3 **b**	4 **y**
5 **g**	6 **y**	*r g b y*	r g b y
7 **b**	r g b y	r g b y	r g b y
r g b y	r g b y	r g b y	r g b y

Figure 2: *Current state of search.*

At this point in the search, seven of the cells have been instantiated (displayed in **boldface** in Figure 2), and the initial domains of the remaining cells are shown. The next cell to instantiate has the domain values in *italics*.

Re-execute (from scratch) the same 7 first assignments in the specified order and, at <u>*each*</u> assignment, draw the Latin Square while filtering the domains of the relative future variables. Do this process for the two following look-ahead strategies:

- the partial look-ahead strategy, forward checking (FC) (5 points)
- the full look-ahead strategy, maintaining arc consistency. (5 points)

Indicate eliminated values by crossing them out or just erasing them.

6 Simple scheduling problem

(Total 10 points)

Courtesy of Rina Dechter

Consider the problem of scheduling five tasks: T_1, T_2, T_3, T_4, and T_5, each of which takes one hour to complete. The tasks may start at 1:00, 2:00, 3:00. Any number of tasks can be executed simultaneously provided the following restrictions are satisfied.

- T_1 must start after T_3.

- T_3 must start before T_4 and after T_5.
- T_2 cannot execute at the same time as T_1.
- T_2 cannot execute at the same time as T_4.
- T_4 cannot start at 2:00.

1. Formulate the problem as a CSP by stating: the variables, their domain, and the applicable constraints. (4 points)
   Hints: focus on the start time of a task.
2. Draw the constraint graph. (2 points)
3. Apply arc-consistency to each constraint in the CSP until no values can be ruled out (i.e., the CSP becomes arc-consistent). (4 points)