CSC384 : Final Project

Christian Lessard

Robertson McClure

Jeffery McBride

Table of Contents

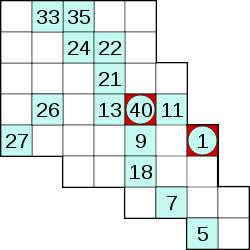
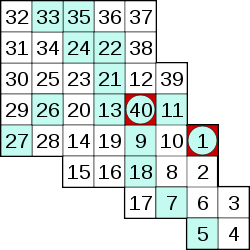
1. Project Motivation
2. Methods
   1. Formulating the CSP – A Change in Implementation
   2. Representing Assignments and Defining Constraints
   3. Comparing Assignment Heuristics
3. Evaluation and Results
4. Limitations and Obstacles
   1. NP-Completeness
   2. Backend Implementation
5. Conclusions
6. Works Cited

01. Project Motivation

*Hidato* is a logic puzzle game invented by Israeli mathematician Dr. Gyora M. Benedek. The game takes the form of a grid hosting several numbers: a *starting number* (typically one), an *ending number* (matching the number of cells on the grid) and some *intermediate numbers* (to limit the number of possible solutions). A player must solve the puzzle by forming a path of adjacencies whereby any given number on the board has both the precedent and subsequent number as a neighbour. Neighbours may be along the horizontal, vertical or diagonal axis.

Hitado is known to be NP-Complete[[1]](#footnote-1), meaning that no known polynomial time algorithm exists to solve instances of the game. Pursuing such would be extremely difficult.

We decided to interpret Hidato as a Constraint Satisfaction Problem (CSP), using techniques from the field of Artificial Intelligence to derive solutions to game instances in a reasonably short amount of time. Such an interpretation also allowed us to test various heuristics and propagators so that we may examine their comparative performance.



*Figure 1: A Hidato board with starting number 1, ending number 40, and intermediate numbers 5, 7, 9, 13, 18, 21, 22, 24, 26, 27, 33, and 35. Image courtesy of Wikipedia.*

*Figure 2: the solution for Figure 1. Image courtesy of Wikipedia.*

02. Methods

**Formulating the CSP – A change in implementation**

After studying CSPs in CSC384, we implemented one as an assignment for a game called Tenner. This assignment used provided backend code that was extremely expensive in terms of space complexity. This was due to the fact that it relied on each constraint having access to a pre-calculated list of every possible satisfying assignment of the variables in its scope. Although this made constraint evaluation quick (with the use of a dictionary), it was taxing on memory and came at the expense of spending time generating each satisfying assignment. We decided that we could rewrite the backend code to dramatically improve this space-complexity issue, without doing so at the cost of efficient time complexity. At first, we believed it would suffice to pass an evaluation function to each constraint, rather than a lookup table of valid solutions. We noticed, however, that constraints in Hitado should only involve pairs of neighbouring numbers. We decided to give every constraint the same structure (checking that three adjacent numbers were placed next to each other on the board), allowing us to define a single, constant time evaluation function for all of them. To check a constraint over three variables, dubbed a “start”, “middle” and “end” variable, one need only apply the following function:

define **check**(constraint):

return next\_to(constraint.start, constraint.middle) and next\_to(constraint.middle, constraint.end)

define **next**\_to(coord1, coord2):

return abs(coord1[0] - coord2[0]) <= 1 and abs(coord1[1] - coord2[1]) <= 1 and not coord1 == coord2

**Representing Assignments** **and** **Defining Constraints**

As previously mentioned, all constraints share the same structure: they ascertain that three adjacent numbers (such as 1, 2 and 3) are placed next to each other on the board. This would have been quite cumbersome had we not made an important decision about encoding data into the CSP’s variables. Initially, we believed it would be sensible to encode each cell on the board as a variable, using possible numeric assignments as the constituents of their current domains. Instead, we decided it would be better to encode the opposite: numbers as variables, with possible coordinates on the board constituting their current domains. Had we gone with the former idea, it would have been difficult to come generalize the structure of our constraints and, more importantly, to create a single, uniform evaluation function for them.

**Comparing Assignment Heuristics**

Another result of our variable encoding what that of the assignment heuristics made available to us. The use of propagators like Forward Checking and Generalized Arc Consistency provide Backtracking Search with a heuristic – known as the Minimum Remaining Value heuristic – for making decisions about which variables to assign next. This heuristic sees Backtracking Search pick whichever variable has the smallest current domain as the next to be assigned. Because we assign coordinates to variables, we were also able to use the “Next In Line” heuristic, which has Backtracking Search pick the next smallest unassigned number. The comparison of these heuristics will be presented in the next section.

03. Evaluation and Results

Here, describe your evaluation objectives and strategy, and your results. In particular, describe the way you’ve chosen to evaluate your approach (i.e. how you will determine if your approach works). Evaluation metrics could include the number of nodes expanded in a search algorithm or the amount of time or memory that you used. We encourage the use of diagrams, graphs, and/or tables to summarize experimental results and to convey important points. Note that it’s ok if your system proves to be inefficient in some way; that’s still a result and we want to know. In addition to graphs and tables, provide a written summary of your findings and their implications, if any.

To evaluate a solution, we wrote a polynomial time verifier function. This function, written as a method called “verify” in class CSP, begins with the variable encoding the starting number 1, and attempts to traverse the subsequent neighbours checking for their adjacencies on the board.

04. Limitations and Obstacles

**NP Completeness**

As mentioned at the beginning of the report, Hitado is proven to be in the class of NP Complete problems. This meant that we would likely be restricted to relatively small boards – at least, smaller than our liking – and that interesting results would come from two particular sources: the implementation of the backend code supporting our CSP, and the comparison of propagators and heuristics applied to Backtracking Search. Despite possible board-dimension limitations, we were correct in believing that interesting discoveries could still be made.

**Backend Implementation**

The most time consuming and difficult part of the project involved writing the base code for the CSP. It was the most bug-prone, and required us to think through everything quite carefully. We’re happy with how things turned out, as we believe our code to be quite fast and elegant. We know that boards of dimensions exceeding 12x12 are too demanding of our CSP to see results in a reasonable amount of time, and it is not immediately obvious how we might circumvent this issue. We also struggled to find more propagators and heuristics to test with our code, simply because many of them are described in research papers we did not have time to read.

05. Conclusions

**Overall Conclusions**

We learned that a well-planned CSP implementation will bare an enormous impact on the results of the problem. We also discovered a game-specific heuristic that lead to interesting findings – namely, a heuristic that actually out did ones discussed in class. It would be nice to explore more propagators for the problem, although this would require more time than we had at our disposal.

**Evaluating Board Difficulty**

It would have also been nice to spend more time developing a stricter measure of difficulty for Hidato boards in general, and testing our work with that measure in mind. Although we did not have time to elaborate on scuh a definition, we did come up with the following ideas that may one day help us develop something more concrete:

1. Board dimensions or, more generally, the number of cells on the boards, are a general indicator of difficulty, because it determines the amount of work to be done.
2. The number of given values on a board reduces its difficulty, as this reduces the total number of cells for which a value must be calculated.
3. Which values are given can have a large impact on difficulty; we found boards of identical dimensions and varying difficulty.

believe that difficulty could be represented as a ratio between board size and the number of given variables on the board, although this would take more thought. Thankfully, our code is modular enough that it should be easy to extend in the future.

06. Works Cited

1. See <http://www.nearly42.org/cstheory/hidato-is-np-complete/> [↑](#footnote-ref-1)