

Explanation Assignment 4

Clue File information extraction

By reading the clue file following information will be extracted. The shape of the puzzle (hexagon or rectangle), the size of the puzzle (size of the hexagon or number of rows and columns), the number of colors and the hexes of the colors and finally all given clues.

Mask generation

For each clue masks are generated, where each mask corresponds to a possibility of the clue in the nonogram. For example, if the nonogram is rectangular and has three rows, four columns, one color and the clue is "2a". Following masks will be generated: "aaBB", "aBBa" and "BBaa", where "B" means the cell should get the background color and "a" means the cell should get the first color. To generate these masks, I wrote a cascading recursive function, which adds a "B" to the mask or uses a clue and tracks the size of the mask. If a clue is used, and the clue is a letter, the letter will be added to the mask n-times, where n is the number in the clue. If the next clue, if it exists, is the same letter an additional "B" is added. "?" clues are tricky. They need a recursive call for each foreground color, but there are multiple restrictions. First if the clue before was a letter and no "B" has been added in between, the "?" can't be the same letter. Second if the clue before was also a "?" and no "B" has been added in between, the current "?" can't be the same letter as the "?" before. Third if after a "?" clue there is a letter clue, a "B" must be added after the guess of the "?" if the guess is the same letter as the letter clue. There are two abort criteria. First if all clues have been used, a valid mask has been found and gets padded with "B"s to match the size of the row/column/hex row. Second if not all clues have been used but the mask is too big, it gets discarded.

Coordinate Systems and color variables

After the generation of the masks, I calculated a coordinate system/ python dictionary for the color variables, which maps each coordinate of a cell in the nonogram to its possible foreground colors. The variables are sequential numbers, one number for each possible foreground color. For example, cell (2,3) -> (13,14) if the nonogram is rectangular and more than one row and four columns. The first coordinate of the hexagonal nonogram is (1, size, 2 * size - 1) where size is the size of the hexagon. In the hexagon there are six different directions where one coordinate stays the same, one is decremented by one and one is incremented by one. For example moving right in the hexagon (0, +1, -1).

DNF, CNF and the Tseytin transformation

By using all masks and the variables of the clues, determined by the row/column/hex row number to get the coordinates which are mapped to the variables of a clue, a DNF can be created. The letters in the mask define if the variables must be negated. For example, a "B" means, all variables for the current coordinate must be negative in the DNF. The DNF subsequently transformed into a CNF by the tseytin transformation for the Sat file. To save computing time the formula of the Tseytin transformation is exploited, so the DNF has not to be calculated. Each possible mask is a clause in the DNF. Every clause leads to an auxiliary variable. The clauses in the CNF are a clause containing all auxiliary variables and a pair of the auxiliary of the current mask together the current variable for each variable for the coordinates of the current mask. I got the idea from https://www.kbyte.io/projects/201908_nonogram/.

Drawing the solution

For drawing the solution, the coordinate systems get reused. By using the solution to determine values of the variables (negative or positive) together with the coordinates of the cells of the nonogram, the cells can be dyed. For example, if all variables of the cell are negative it gets dyed in the background color. If the second variable is positive, it gets dyed in the second foreground color. The coordinates are also used to calculate the position in the Svg.

Uniqueness

By adding one clause containing all variables, without the auxiliary variables, together with the status (negative or positive) in the solution, this solution will be banned. If the Sat-Solver determines that the extended CNF is still satisfiable the puzzle isn't unique.