Abstraction for Zooming-In to Unsolvability Reasons of Grid-Cell Problems ⋆ (Extended Abstract)

Thomas Eiter, Zeynep G. Saribatur, Peter Schüller

Institute of Logic and Computation, TU Wien

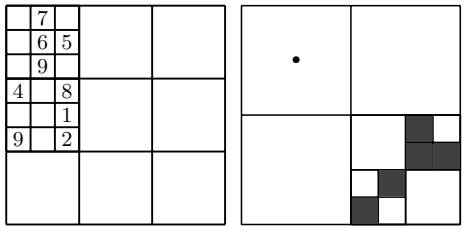
Abstraction is about focusing on the relevant details and disregarding the irrelevant ones that are not really needed to be taken into account. Humans are capable of abstracting away irrelevant details when studying problems, especially to point out the details that cause a problem to be not solvable and provide explanations. In graph coloring, for instance, if a given graph is non-colorable, finding some subgraph (e.g., a clique) of it which causes the unsolvability, and not caring about other nodes, is a typical abstraction a human would do. The abstraction capability is especially noticeable for problems over grid-cells, as humans are able to disregard certain parts of the grid and focus on the key elements important for the problem. Empowering the machine with such an abstraction capability to obtain human-like machine explanations is one of the challenges of explainable AI.

Answer set programming (ASP) [1], with its expressivity and representation power, is a convenient tool for investigating ways of applying human-inspired problem solving methods. Ongoing studies in understanding how ASP programs find a solution (or none) to a problem mainly focus on debugging answer sets or finding justifications (see the recent survey [2]). However, as noted in [2], the obtained explanations may contain too many details which prevent one from seeing the crucial parts. This is where some notion of abstraction would come in handy. Recently, the notion of abstraction has been introduced for ASP [3], by means of clustering the elements of the domain and automatically constructing an over-approximation of a given program. Employing such an abstraction showed potential for aiding program analysis as it allows for problem solving over abstract notions, by achieving concrete abstract answer sets that reflect relevant details only.

Problems that involve multi-dimensional structures, e.g. grid-cells, require a differentiated view of an abstraction in order to provide insight that is similar to humans, by focusing on certain areas and abstracting away the rest. Sudoku is a well-known problem, where the empty cells need to be filled with the guidance of the given numbers by respecting some constraints. Fig. 1(a) shows an instance with the focus on the sub-regions that contain the reason why a solution can not be found. Since the numbers 6,7 appear in the middle column, they can only be assigned to the below region's left column, which is not possible as only one empty cell exists. As another example, Fig. 1(b) shows an instance for the

⋆ Appears in the Proceedings of the IJCAI 2019 Workshop on Explainable Artificial Intelligence.

Fig. 1. Unsatisfiable abstractions of grid-cell problems



(a) Sudoku (b) Reachability

Reachability problem where some cells are not reachable from the upper-left corner due to the obstacles in the focused area.

In this paper, we empower the approach in [3] to handle such a hierarchical view of abstraction that automatically adjusts the granularity towards the relevant details for the problem. The method is used for *zooming in* to the area that is sufficient for the machine to realize the unsolvability of a problem instance. Distinguishing this area becomes the machine's way of explaining unsolvability, which is then compared with how humans provide explanations.

Our contributions are briefly summarized as follows:

- We introduce multi-dimensional abstraction mappings over a domain. For this, we ought to modify the previous abstraction method [3] by having an existential abstraction over the relations, in order to enable dealing with elements of different abstraction layers.
- We extend the abstraction-&-refinement methodology with handling the structural aspects of grid-cells by using a quad-tree abstraction and consider more sophisticated decision making approaches on the refinement to observe its effects on the resulting abstractions.
- We use this approach in detecting the unsolvability of benchmarks problems involving grid-cells. A user study is conducted to compare the resulting abstractions with human explanations, which showed that such a hierarchic abstraction can provide intuitive and "to the point" explanations of unsolvability. The user study on human explanations also revealed the implicit abstraction capabilities of humans and the acknowledged need for studying the meaning of explanation.

References

1. Erdem, E., Gelfond, M., Leone, N.: Applications of answer set programming. AI Magazine **37**(3), 53–68 (2016)
2. Fandinno, J., Schulz, C.: Answering the "why" in answer set programming - A survey of explanation approaches. TPLP **19**(2), 114–203 (2019)
3. Saribatur, Z.G., Schüller, P., Eiter, T.: Abstraction for non-ground answer set programs. In: Proc. of JELIA, pp. 576–592. LNCS, Springer (2019)