Centre for
Data Analytics

Insight

Solving Logic Grid Puzzles with an Algorithm that Imitates Human Behavior

http://ucc.insight-centre.org/gescamocher/Grail.zip

Guillaume Escamocher Barry O'Sullivan

September 30, 2019

DCU NUI Galway OÉ Gaillimh UCC University College Cork, Ireland Coláiste na hOllscoile Corcaigh UCD Dublin SFI Science Foundation Ireland

Formal Definition

Logic grid puzzle

- k categories
- n elements in each category
- clues

Solution

A matching of the elements into k-tuples respecting all clues.
A valid logic grid puzzle has exactly one solution.

Running example

Categories

- First Name: “Angela”, “Donald”, “Leo”
- Country: “Germany”, “Ireland”, “United States”
- Year of Birth: “1946”, “1954”, “1979”

Clues

1. The person from the “United States” was born in “1946”.
2. “Leo” is younger than the person from “Germany”.
3. “Donald” was born in “1946” or he is from “Ireland”.

Solution

- ⟨ “Angela”, “Germany”, “1954”⟩
- ⟨ “Donald”, “United States”, “1946”⟩
- ⟨ “Leo”, “Ireland”, “1979”⟩

Outline

Step 1: Acquisition (partial)

- User is asked to enter the puzzle.
- Clue $\Leftrightarrow$ Set of constraints
- For each constraint, the user picks among the options offered by the program.
- Input files for the website puzzles can be found at http://ucc.insight-centre.org/gescamocher/Grail_Input.zip

Step 2: Modelling

Grid

Each pair of elements from different categories is represented by a cell. Solving a puzzle is filling each cell with *yes* if the two elements are to be matched, or with *no* otherwise.

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela						
	Donald						
	Leo						
Year of Birth	1946						
	1954						
	1979						

Our implementation

- Array of integers.
- Each cell is initialized at 0.
- Value set to 1 (resp. -1) when cell determined to be *yes* (resp. *no*).

Step 3: Solving

One rule, one cell

At each step, the algorithm uses one inference rule to fill one cell of the grid.

Goal: to reason like a human

- same inference rules
- same order

Inference Rules

- Clues rules (58): information that can be derived from the clues.
- Consistency rules: information that can be derived from just the current state of the grid.
  - Basic consistency rules (2): easy, humans love them. They correspond to the *Arc Consistency* and *Clique Consistency* inference rules[1].
  - Advanced consistency rules (3): tedious, humans try to avoid them. Two of them correspond to the *Transitivity* and *Generalized Neighborhood Clique Consistency* inference rules[1].

[1] M. H. Sqalli and E. C. Freuder, *Inference-Based Constraint Satisfaction Supports Explanation*, AAAI 1996

Priority

1. Cycle through the clues in order. For each one, check whether a clues rule can be applied. No going back to previous clues, even if progress is made.
2. Apply basic consistency rules as much as possible. If progress is made, check if basic consistency rules that failed previously now work.
3. Only try advanced consistency rules if no progress was made with other types of rules. As soon as any advanced consistency rule fills a cell, go to 1.

Step 4: Explanation

One cell, one line

Every time a cell is filled, a line of explanation is written.

Output files for the website puzzles can be found at
http://ucc.insight-centre.org/gescamocher/Grail_Output.zip

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela						
	Donald						
	Leo						
Year of Birth	1946						
	1954						
	1979						

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela						
	Donald						
	Leo						
Year of Birth	1946			Y			
	1954						
	1979						

- "United States" is "1946" (Clue 1).

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela						
	Donald						
	Leo	•					
Year of Birth	1946			Y			
	1954						
	1979						

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

- "United States" is "1946" (Clue 1).
- "Leo" is after "Germany" in the "Year of Birth" category (Clue 2), so "Leo" is not "Germany".

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela						
	Donald						
	Leo	•			•		
Year of Birth	1946			Y			
	1954						
	1979						

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

- "United States" is "1946" (Clue 1).
- "Leo" is after "Germany" in the "Year of Birth" category (Clue 2), so "Leo" is not "Germany".
- "Leo" is after "Germany" in the "Year of Birth" category (Clue 2), so "Leo" is not the first element in that category, so "Leo" is not "1946".

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela						
	Donald						
	Leo	•			•		
Year of Birth	1946			Y			
	1954						
	1979	•					

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

- "United States" is "1946" (Clue 1).
- "Leo" is after "Germany" in the "Year of Birth" category (Clue 2), so "Leo" is not "Germany".
- "Leo" is after "Germany" in the "Year of Birth" category (Clue 2), so "Leo" is not the first element in that category, so "Leo" is not "1946".
- "Leo" is after "Germany" in the "Year of Birth" category (Clue 2), so "Germany" is not the last element in that category, so "Germany" is not "1979".

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela						
	Donald						
	Leo	•			•		
Year of Birth	1946	•	•	Y			
	1954	Y	•	•			
	1979	•	Y	•			

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

- 7 cells can be filled from basic consistency.

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela						
	Donald						
	Leo	•			•	•	
Year of Birth	1946	•	•	Y			
	1954	Y	•	•			
	1979	•	Y	•			

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

- "Germany" is not one of the first 1 element in the "Year of Birth" category, and "Leo" is after "Germany" in that category (Clue 2), so "Leo" is not one of the first 2 elements in the "Year of Birth" category, so "Leo" is not "1954".

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela						
	Donald	•					
	Leo	•			•	•	
Year of Birth	1946	•	•	Y			
	1954	Y	•	•			
	1979	•	Y	•			

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

- "Germany" is not one of the first 1 element in the "Year of Birth" category, and "Leo" is after "Germany" in that category (Clue 2), so "Leo" is not one of the first 2 elements in the "Year of Birth" category, so "Leo" is not "1954".
- "Donald" is "1946" or "Ireland" (Clue 3), and "Germany" is neither "1946" nor "Ireland", so "Donald" is not "Germany".

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela						
	Donald	•				•	
	Leo	•			•	•	
Year of Birth	1946	•	•	Y			
	1954	Y	•	•			
	1979	•	Y	•			

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

- "Germany" is not one of the first 1 element in the "Year of Birth" category, and "Leo" is after "Germany" in that category (Clue 2), so "Leo" is not one of the first 2 elements in the "Year of Birth" category, so "Leo" is not "1954".
- "Donald" is "1946" or "Ireland" (Clue 3), and "Germany" is neither "1946" nor "Ireland", so "Donald" is not "Germany".
- "Donald" is "1946" or "Ireland" (Clue 3), and "1954" is neither "1946" nor "Ireland", so "Donald" is not "1954".

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela	Y	•	•	•	Y	•
	Donald	•			Y	•	•
	Leo	•			•	•	Y
Year of Birth	1946	•	•	Y			
	1954	Y	•	•			
	1979	•	Y	•			

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

- 9 cells can be filled from basic consistency.

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela	Y	•	•	•	Y	•
	Donald	•		Y	Y	•	•
	Leo	•			•	•	Y
Year of Birth	1946	•	•	Y			
	1954	Y	•	•			
	1979	•	Y	•			

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

- 9 cells can be filled from basic consistency.
- "Donald" is "1946" and "1946" is "United States", so "Donald" is "United States".

		Country			Year of Birth		
		Germany	Ireland	USA	1946	1954	1979
First Name	Angela	Y	•	•	•	Y	•
	Donald	•	•	Y	Y	•	•
	Leo	•	Y	•	•	•	Y
Year of Birth	1946	•	•	Y			
	1954	Y	•	•			
	1979	•	Y	•			

Clues

1. The person from the "United States" was born in "1946".
2. "Leo" is younger than the person from "Germany".
3. "Donald" was born in "1946" or he is from "Ireland".

- 9 cells can be filled from basic consistency.
- "Donald" is "1946" and "1946" is "United States", so "Donald" is "United States".
- 3 cells can be filled from basic consistency.

Bonus Features

Clue obsolescence User is notified when a clue will not be looked at again.

Single cell solving Only the cells/explanation leading to the solving of a given cell are filled/displayed.

CNF conversion Puzzles can be written into a CNF file. Useful for debugging.

Outline

Puzzles Considered

Difficulty	Easy	Hard	Zebra	Total
Valid	55	13	1	69
Invalid	5	3		

Most invalid puzzles removed because of too many (> 1) solutions.

Challenges Encountered

- Information in the opening statement is needed.
  - Easy 22 “Baggage Mishaps”
  - Easy 60 “Holiday Decision”
  - Easy 64 “Robbery at Millionaire’s Mansion”
  - Easy 70 “The racehorses”
  - Easy 76 “Three little boys”
  - Easy 83 “The Enchanted Forest”

  Fix: treated opening statement as Clue 0.
- Cross-referencing meta-information that our constraints cannot model.
  - Easy 65 “Sporting Excellence”

  Fix: none found :(
- Objective contains a label that appear in the clues but not in the existing categories.
  - Hard 119 “A New Personal Computer” (label is “Andrew”, objective is “which computer has been chosen by Andrew?”)

  Fix: added an extra category containing the elements {“Andrew”, “NotAndrew1”, “NotAndrew2”, “NotAndrew3”, “NotAndrew4”}.

Success Rate

68/69 (98.6%)

How long does it take to process all these puzzles?

Conclusion

Our program

- Fulfills 3.5/4 steps of the challenge.
- Solved with explanation 67/68 puzzles from the target website, including all the valid Hard ones.
- Runs quickly, anywhere.
- Contains additional features to enhance explanation.

The End

Thank you for your attention
Questions are welcome!

- http://ucc.insight-centre.org/gescamocher/Grail.zip
- http://ucc.insight-centre.org/gescamocher/Grail_Input.zip
- http://ucc.insight-centre.org/gescamocher/Grail_Output.zip