# Information and coding theory

#### Project I

## **Practical informations**

Each project should be executed by groups of two students. We expect each group to provide:

- A brief report (in PDF format) collecting the answers to the different questions.
- The scripts you have implemented.

The report and the scripts should be submitted as a tar.gz (or zip) file on Montefiore's submission plateform (http://submit.montefiore.ulg.ac.be) before  $March\ 21,\ 23:59\ GMT+2$ . You must concatenate your sXXXXXX ids as group, archive and report names.

### Questions

#### Information measures

#### Exercises by hand

Let  $\mathcal{X}$  and  $\mathcal{Y}$  be two random variables that take on values  $x_1, x_2, x_3, x_4$  and  $y_1, y_2, y_3, y_4$ . The joint distribution of these two random variables is as follows:

|                     | $x = x_1$ | $x = x_2$ | $x = x_3$ | $x = x_4$ |
|---------------------|-----------|-----------|-----------|-----------|
| $y = y_1$           | 1/8       | 1/16      | 1/16      | 1/4       |
| $y = y_2$           | 1/16      | 1/8       | 1/16      | 0         |
| $y = y_3$           | 1/32      | 1/32      | 1/16      | 0         |
| $y = y_3$ $y = y_4$ | 1/32      | 1/32      | 1/16      | 0         |

Let  $\mathcal{W}$  and  $\mathcal{Z}$  be two binary random variables whose values are determined as follows:

#### Calculate

- 1.  $H(\mathcal{X}), H(\mathcal{Y}), H(\mathcal{W}), H(\mathcal{Z})$
- 2.  $H(\mathcal{X}, \mathcal{Y}), H(\mathcal{X}, \mathcal{W}), H(\mathcal{Y}, \mathcal{W}), H(\mathcal{W}, \mathcal{Z})$
- 3.  $H(\mathcal{X}|\mathcal{Y}), H(\mathcal{W}|\mathcal{X}), H(\mathcal{Z}|\mathcal{W}), H(\mathcal{W}|\mathcal{Z})$
- 4.  $H(\mathcal{X}, \mathcal{Y}|\mathcal{W}), H(\mathcal{W}, \mathcal{Z}|\mathcal{X})$
- 5.  $I(\mathcal{X}; \mathcal{Y}), I(\mathcal{X}; \mathcal{W}), I(\mathcal{Y}; \mathcal{Z}), I(\mathcal{W}; \mathcal{Z})$
- 6.  $I(\mathcal{X}; \mathcal{Y}|\mathcal{W}), I(\mathcal{W}; \mathcal{Z}|\mathcal{X})$

#### **Implementation**

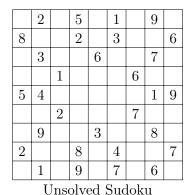
In Python or Julia.

- 7. Write a function *entropy* that computes  $H(\mathcal{X})$ , the entropy of a random variable  $\mathcal{X}$ , given<sup>1</sup> its probability distribution  $P_{\mathcal{X}} = (p_1, p_2, \ldots, p_n)$ . What are the key parts of your implementation? Intuitively, what is measured by the entropy?
- 8. Let  $\mathcal{X}$  and  $\mathcal{Y}$  be two discrete random variables. Write a function  $joint\_entropy$  that computes  $H(\mathcal{X}, \mathcal{Y})$ , the joint entropy of  $\mathcal{X}$  and  $\mathcal{Y}$ . What are the key parts of your implementation? Compare this function with the entropy function, what do you notice?
- 9. Let  $\mathcal{X}$  and  $\mathcal{Y}$  be two discrete random variables. Write a function  $conditional\_entropy$  that computes  $H(\mathcal{X}|\mathcal{Y})$ , the conditional entropy of  $\mathcal{X}$  given  $\mathcal{Y}$ . What are the key parts of your implementation?
- 10. Let  $\mathcal{X}$  and  $\mathcal{Y}$  be two discrete random variables. Write a function  $mutual\_information$  that computes  $I(\mathcal{X}; \mathcal{Y})$ , the mutual information between  $\mathcal{X}$  and  $\mathcal{Y}$ . What are the key parts of your implementation? What can you deduce from the influence of one variable on the other?
- 11. Let  $\mathcal{X}$ ,  $\mathcal{Y}$  and  $\mathcal{Z}$  be three discrete random variables. Write functions  $cond\_joint\_-$  entropy and  $cond\_mutual\_information$  that compute  $H(\mathcal{X}, \mathcal{Y}|\mathcal{Z})$  and  $I(\mathcal{X}; \mathcal{Y}|\mathcal{Z})$  respectively. To do so, extend the  $joint\_entropy$  and  $mutual\_information$  functions.

#### Computer-aided exercises

12. Using implemented functions, verify and compare your results of questions 1 to 6.

# Designing informative experiments



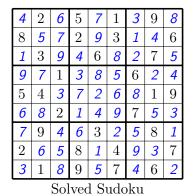


Figure 1: Example of a Sudoku grid and its solution.

The goal of this logic-based game is to fill in a  $9 \times 9$  grid with digits so that each column, each row, and each of the nine  $3 \times 3$  subgrids that compose the grids contain all of the digits from 1 to 9. Typically, an *unsolved Sudoku* (see left grid on Figure 1) is a partially completed grid where each given digit is called a *clue*.

Let us associate a random variable  $\mathcal{X}$  to each square of the grid ( $\mathcal{X}_1$  to  $\mathcal{X}_9$  on first row,  $\mathcal{X}_{10}$  to  $\mathcal{X}_{18}$  on second row, etc. until  $\mathcal{X}_{81}$ ). The entropy of the grid is the sum of the entropies of all random variables  $\mathcal{X}_1, \ldots, \mathcal{X}_{81}$ .

<sup>&</sup>lt;sup>1</sup>In practice, you will pass the appropriate probability distribution(s) as argument(s) of the functions.

- 13. What is the entropy of a single square (i.e., a random variable) independently of others?
- 14. What is the entropy of the following subgrid?

|   | 2 |  |
|---|---|--|
| 8 |   |  |
|   | 3 |  |

- 15. What is the entropy of the unsolved Sudoku grid?
- 16. Using information theory, how would you proceed to solve the Sudoku? Which squares would you fill in first? Justify.
- 17. Let us assume that you can choose *one* additional clue (i.e., revealing the correct digit in an empty square). Which one would you choose and why?
- 18. Let us assume that you can choose *sequentially* more than one clue. Design a strategy (using information theory) that determines the next clue to reveal.
- 19. Let us assume that you can choose *simultaneously* more than one clue. Design a strategy (using information theory) that determines the next clues to reveal (at once).