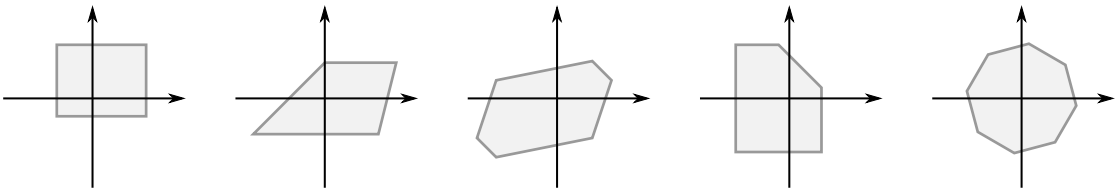
Exercise 06

Zonotopes and Abstract Interpretation

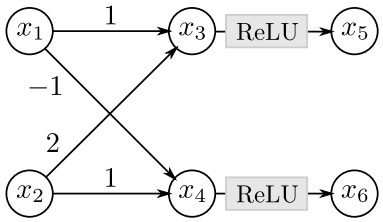
Reliable and Interpretable Artificial Intelligence
ETH Zurich

Problem 1 (Zonotope Concretizations)**.** Which of the following 2D regions (a–e) represent concretizations of a zonotope? For all such regions, sketch a set of 2D magnitude vectors $a_0, \ldots, a_k$ describing the zonotope (select k as small as possible).



(a) (b) (c) (d) (e)

Problem 2 (Certification using Zonotopes)**.** Consider the following small neural network with two input neurons x_1, x_2 and two output neurons x_5, x_6. The network consists of an affine layer followed by a ReLU layer.



$$x_5 = \mathrm{ReLU}(x_1 + 2x_2)$$
$$x_6 = \mathrm{ReLU}(x_2 - x_1)$$

You are given the following zonotope ψ over the input neurons:

$$\psi: \quad \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \cdot \varepsilon_1 + \begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \varepsilon_2 + \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

Your goal is to prove that $x_5 \geq x_6$ for all inputs x_1, x_2 in the zonotope ψ.

1. Draw the concretization $\gamma(\psi)$ of ψ. What shape does it have?

2. Using the transformers for affine and ReLU layers discussed in the lecture, transform ψ to a zonotope ϕ over the output neurons of the network above.

3. Draw the concretization $\gamma(\phi)$ of ϕ. Can you use ϕ to prove the desired property?

Problem 3 (hybrid Zonotopes)**.** In this problem, we consider a new convex relaxation, fusing the zonotope with the interval relaxation. Specifically, we generalize the standard zonotope $x = c + \sum_i a_i \epsilon_i$ for $c \in \mathbb{R}$ and $a_i \in \mathbb{R}$ for all i to $x = [c_l, c_u] + \sum_i a_i \epsilon_i$, where we replaced the center c with an interval $[c_l, c_u]$. The design goal for this exercise is to not increase the number of error-terms.

1. Derive a formula for the addition of two hybrid Zonotopes.

2. Derive a formula for the unary "-" operation applied to a hybrid Zonotope.

3. Derive a formula for the subtracting one hybrid Zonotope from another one.

4. Derive a formula for concretizing a hybrid Zonotope to an interval.

5. Derive a formula for multiplying an interval with a hybrid Zonotope.

Bonus: Derive a formula for the multiplication of two hybrid Zonotopes.