

# GEDLIB: A C++ Library for Graph Edit Distance Computation

D. B. Blumenthal<sup>1, 2, ✉</sup>, N. Boria<sup>3</sup>, S. Bougleux<sup>3</sup>, L. Brun<sup>3</sup>, J. Gamper<sup>1</sup>

<sup>1</sup>Free University of Bozen-Bolzano, Faculty of Computer Science, Bolzano, Italy

<sup>2</sup>Technical University of Munich, Chair of Experimental Bioinformatics, Freising, Germany

<sup>3</sup>Normandie Université, UNICAEN, ENSICAEN, CNRS, GREYC, Caen, France

✉david.blumenthal@wzw.tum.de, david.blumenthal@inf.unibz.it

<https://github.com/dbblumenthal/gedlib>

## GRAPH EDIT DISTANCE (DEFINITION)

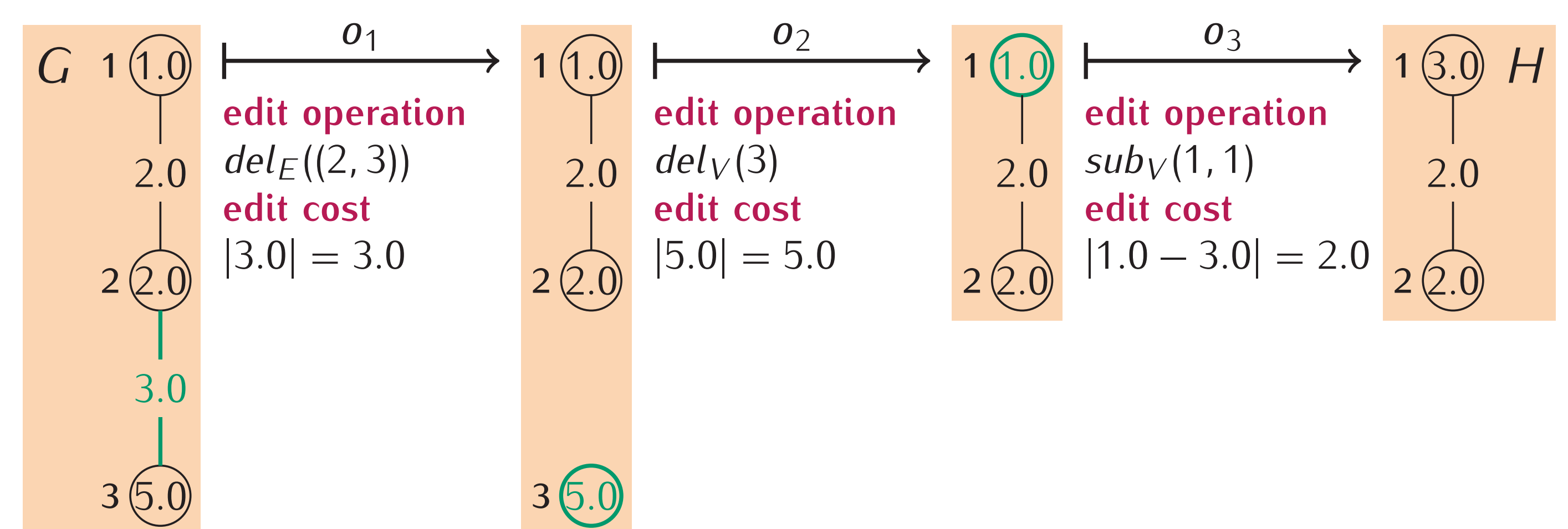
- **GED( $G, H$ )** :=  $\min\{c(P) \mid P \text{ edit path between } G \text{ and } H\}$
- **labeled graphs**:  $G = (V^G, E^G, \ell_V^G, \ell_E^G)$  and  $H = (V^H, E^H, \ell_V^H, \ell_E^H)$
- **labeling functions**:  $\ell_V^G : V^G \rightarrow \Sigma_V$  and  $\ell_E^G : E^G \rightarrow \Sigma_E$
- **node and edge label spaces**:  $\Sigma_V$  and  $\Sigma_E$
- **edit path**: edit operation sequence  $P = (o_1, \dots, o_r)$  transforms  $G$  into  $H$
- **edit operations and edit costs** ( $i \in V^G, k \in V^H, (i, j) \in V^G, (k, l) \in E^H$ ):

node edit op.	edit cost	edge edit op.	edit cost
$sub_V(i, k)$	$c_V(\ell_V^G(i), \ell_V^H(k))$	$sub_E((i, j), (k, l))$	$c_E(\ell_E^G(i, j), \ell_E^H(k, l))$
$del_V(i)$	$c_V(\ell_V^G(i), \epsilon)$	$del_E((i, j))$	$c_E(\ell_E^G(i, j), \epsilon)$
$ins_V(k)$	$c_V(\epsilon, \ell_V^H(k))$	$ins_E((k, l))$	$c_E(\epsilon, \ell_E^H(k, l))$

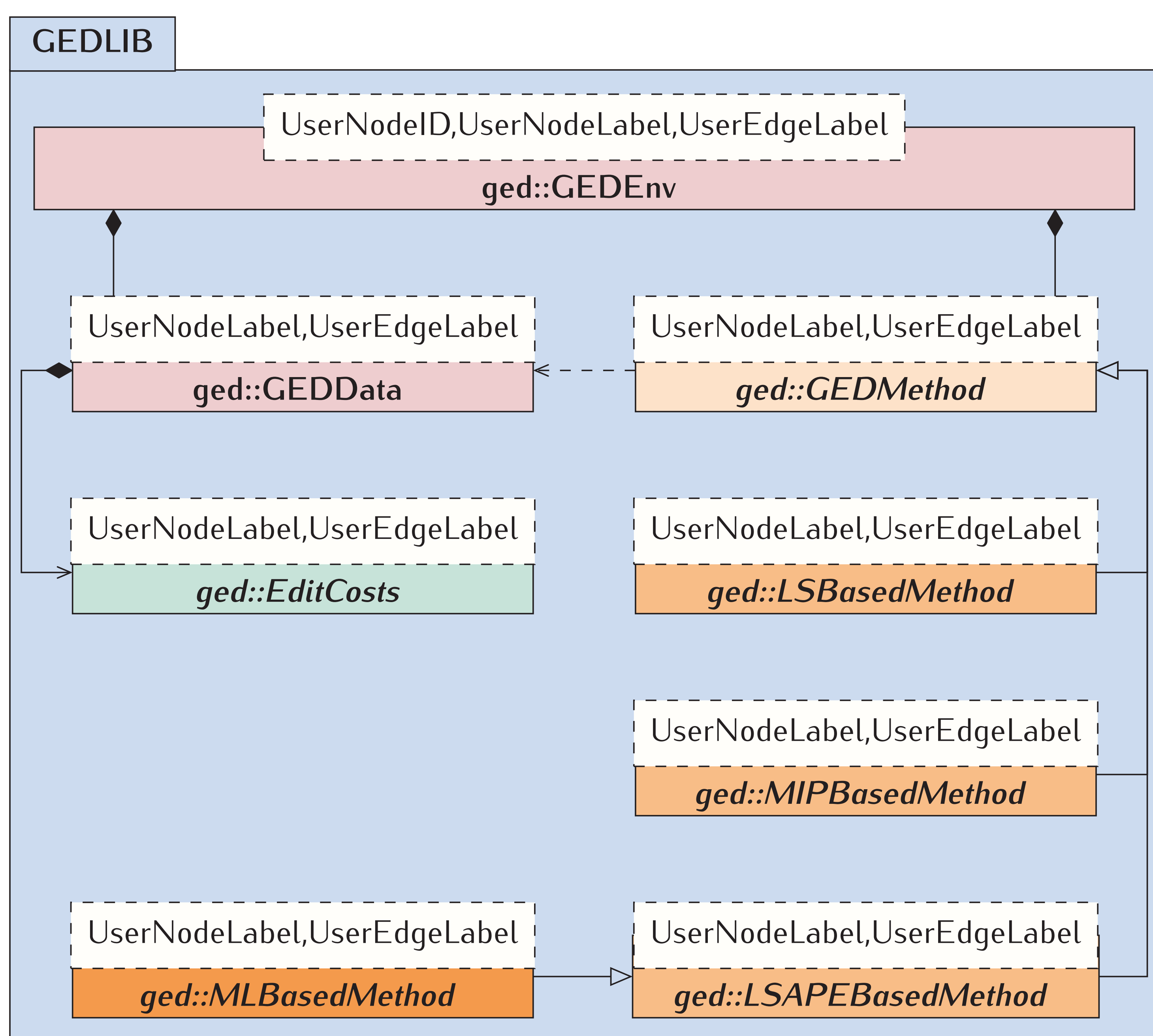
- **edit path cost**:  $c(P) = \sum_{i=1}^r c(o_i)$

## GRAPH EDIT DISTANCE (EXAMPLE)

- **node and edge label spaces**:  $\Sigma_V = \Sigma_E = \mathbb{R}_{\geq 0}$
- **sub. costs**:  $c_V(\alpha, \beta) = c_E(\alpha, \beta) = |\alpha - \beta|$
- **del. and ins. costs**:  $c_V(\alpha, \epsilon) = c_V(\epsilon, \alpha) = c_E(\alpha, \epsilon) = c_E(\epsilon, \alpha) = |\alpha|$
- **edit path cost**:  $c(P) = c(o_1) + c(o_2) + c(o_3) = 3.0 + 5.0 + 2.0 = 10.0$



## ARCHITECTURE



## USER INTERFACE

UserNodeID, UserNodeLabel, UserEdgeLabel	
ged::GEDEnv	
...	// misc. variables
+ load_gxl_graph()	// loads graph given as GXL file
+ load_exchange_graph()	// loads graph from memory
+ set_edit_costs()	// selects the edit costs
+ init()	// initializes the environment
+ set_method()	// selects the GED method
+ run_method()	// runs the selected GED method
...	// misc. member functions

## ABSTRACT CLASS FOR IMPLEMENTING EDIT COSTS

UserNodeLabel, UserEdgeLabel	
ged::EditCosts	
+ node_del_cost_fun()	// computes node deletion cost
+ node_ins_cost_fun()	// computes node insertion cost
+ node_rel_cost_fun()	// computes node relabelling cost
+ edge_del_cost_fun()	// computes edge deletion cost
+ edge_ins_cost_fun()	// computes edge insertion cost
+ edge_rel_cost_fun()	// computes edge relabelling cost
...	// misc. member functions

## FEATURES

- **use case**: compute distances between undirected graphs
- **algorithms**: over 30 state-of-the-art GED algorithms available in GEDLIB
- **edit costs**: edit costs for benchmark graphs available in GEDLIB
- **genericity**: easily implement edit costs for your graphs and label spaces
- **extensibility**: easily implement new GED algorithms
- **usability**: installation script and extensive Doxygen documentation

## COMING SOON

- **Python**: Python bindings for even better usability
- **median graphs**: quickly compute generalized median graphs
- **clustering**: cluster the graphs in your dataset
- **data reduction**: reduce your data by computing representatives
- **indexing**: index your graph datasets for fast similarity queries
- **even more algorithms**: provided by the pattern recognition community

## PUBLICATIONS THAT PRESENT OR USE GEDLIB

- [1] D. B. Blumenthal, S. Bougleux, J. Gamper, and L. Brun, "GEDLIB: A C++ library for graph edit distance computation," in *GbRPR*, 2019, pp. 14–24.
- [2] D. B. Blumenthal, N. Boria, J. Gamper, S. Bougleux, and L. Brun, "Comparing heuristics for graph edit distance computation," *VLDB J.*, 2019, in press.
- [3] N. Boria, D. B. Blumenthal, S. Bougleux, and L. Brun, "Improved local search for graph edit distance," *Pattern Recognit. Lett.*, 2019, accepted.
- [4] D. B. Blumenthal, S. Bougleux, J. Gamper, and L. Brun, "Ring based approximation of graph edit distance," in *S+SSPR*, 2018, pp. 293–303.

Presented at ISMB/ECCB 2019, Basel, Switzerland, July 21–25, 2019