An Evolutionary Strategy for Automatic Hypotheses Generation inspired by Abductive Reasoning

Roberto Pietrantuono
Università degli Studi di Napoli Federico II
Napoli, Italy
roberto.pietrantuono@unina.it

ABSTRACT

This paper proposes a new evolutionary strategy – called *Evolutionary Abduction* (EVA) - designed to target a class of problems called Combinatorial Causal Optimization Problems (CCOP). In a CCOP, the goal is to find combinations of causes that best explain or predict an effect of interest. EVA is inspired by abduction, a powerful form of causal inference employed in many artificial intelligence tasks. EVA defines a set of abductive *operators* to repeatedly construct hypothetical cause-effect instances, and then automatically assesses their *plausibility* as well as their *novelty* with respect to already known instances. Experiments confirm that, given a background knowledge, EVA can construct better hypotheses for a given effect, outperforming alternative strategies based on common metaheurstics previously used for CCOP.

CCS CONCEPTS

• Computing methodologies → Discrete space search; **• Mathematics of computing → Combinatorial optimization**.

KEYWORDS

Evolutionary algorithm, Causal reasoning

ACM Reference Format:
Roberto Pietrantuono. 2023. An Evolutionary Strategy for Automatic Hypotheses Generation inspired by Abductive Reasoning. In *Genetic and Evolutionary Computation Conference Companion (GECCO '23 Companion), July 15–19, 2023, Lisbon, Portugal.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3583133.3590568

1 INTRODUCTION

A recently presented class of optimization problems is Combinatorial Causal Optimization Problems (CCOP) [12]. A CCOP formulates the causal inference problem of finding the best explanatory causes for an effect as an optimization task, in which searching for a solution means hypothesizing a suitable set of causes for the effect of interest. The problem has been addressed by customizing conventional metaheuristics with promising results. However, given the nature of the task (i.e., causal inference), this paper takes a different perspective. I hereafter propose a new strategy inspired by abduction, called *Evolutionary Abduction* (EVA).

Abduction is a powerful form of causal reasoning frequently employed in everyday common-sense reasoning and as first step of scientific reasoning [11], [6]. Abductive reasoning infers *possible causes* for a given *effect*, by advancing *hypotheses* or generating new ideas outside the given facts [4]. As such, abduction is said to be an *ampliative* form of inference, as it is able to enlarge our knowledge, but *uncertain*, because its inferences are more susceptible to error than deductive and inductive ones, and need to be "validated".

EVA mimics the process of actively searching for explanations for a given observation, by generating hypotheses for *plausible* causes of an effect, exploiting both an *experience-based* knowledge and the *ontological knowledge* a human has about the phenomena in explanation. Like it happens in human abductive reasoning, it is generally not known whether the hypothseized explanations (i.e,. solutions) are admissible beforehand, but the hypotheses need to be assessed for their plausibility based on background knowledge. Only plausible explanations survive and are proposed as solutions.

EVA defines three operators to mimic and automate the most common *patterns* of abduction [13], which ultimately lead to construct cause-effect combinations as solutions to CCOP. These are then automatically assessed for their plausibility by exploiting the background knowledge stored as an archive of past observations.

EVA is experimented on four real-world datasets, having a number of variables (namely, co-occurring causes for a given effect) that ranges from 9 to 27. Results demonstrate the potential of EVA: using a small fraction (10%) of the datasets as knowledge base, EVA significantly outperforms four alternative metaheuristics previously used for CCOP. The code, datasets, and the Appendix are at: `https://github.com/rpietrantuono/MOEVA/`

2 CAUSAL OPTIMIZATION

Hereafter, I recall the *Combinatorial Causal Optimization Problem* formulation [12]. In a CCOP, the goal is to find a proper set of *causes* A_i (or *explanations* or *hypotheses*) for a given set of *effects* B_j that minimize (maximize) one or more objectives. Causes and effects are abstractions of *phenomena/events* regarding any element of interest $i \in U$, where U is the domain of the problem to be solved (i.e., the set of elements possibly involved in the inference). Specifically: causes and effects are *literals*, namely atomic formulae or their negation (a.k.a. *atoms*). In first-order logic, atoms correspond to predicate symbols together with their arguments, and a cause-effect pair to infer is a *clause* conveniently represented as a *rule* $B_j \leftarrow A_i$.

The literals are the decision variables (DV), x_i ($i = 1, ..., n = |U|$). Each x_i takes values in a non-empty discrete set representing its domain, D_i. For instance, a decision variable x_i in a medical dataset can represent a blood analysis parameter, taking values from a discrete set D_i={M, C}: "*M: moderately over-threshold*", "*C: critically*

GECCO '23 Companion, July 15–19, 2023, Lisbon, Portugal

ACM ISBN 979-8-4007-0120-7/23/07.
https://doi.org/10.1145/3583133.3590568

over-threshold". The set of (discrete) DVs $X = \{x_1, ..., x_n\}$ is the union of two non-empty disjoint subsets related by a *causality* relationship ($\xrightarrow{c}$)[1]: X_s, that is the set of causes (named *sources*), and X_t, the set of effects to be explained (named *targets*): $X = X_s \cup X_t = \{x_{s_1}, ..., x_{s_j}; x_{t_{j+1}}, ..., x_{t_n}\}$, and: $\forall x_s \in X_s, \exists x_t \in X_t : x_s \xrightarrow{c} x_t$.

Each source (target) variable $\{x_{s_1}, \dots, x_{s_j}\}$ ($\{x_{t_{j+1}}, \dots, x_{t_n}\}$) takes values in the respective discrete set: $D_s = \{D_{s_1}, \dots D_{s_j}\}$ ($D_t = \{D_{t_{j+1}}, \dots D_{t_n}\}$). A CCOP is expressed as follows:

$$\begin{aligned} \text{Maximize} \quad & \pi(\mathbf{x}), \nu(\mathbf{x}) \\ & \mathbf{C} = (\mathbf{C}_k, \mathbf{C}_u) = (c_{k_1}, \dots, c_{k_q}; c_{u_{q+1}}, \dots, c_{u_l}) \\ & \mathbf{x} = (\mathbf{x}_s; \mathbf{x}_t) = (x_{s_1}, \dots, x_{s_j}; x_{t_{j+1}}, \dots, x_{t_n}) \in \Omega \quad (1) \end{aligned}$$

- $\Omega = \{D_s \cup D_t\}^n$ is the decision space, the set of all possible values that decision variables can take. In the abduction metaphor, Ω represents the ontological knowledge (or *ontology*) of the domain, namely all the causes and effects that can concur to the inference.
- $\mathbf{x}$ is a candidate solution. A solution proposes an explanation for the effect(s) in $\mathbf{x}_t$ by potential cause(s) in $\mathbf{x}_s$ (with $1 \le |\mathbf{x}_s|, |\mathbf{x}_t| \le n$). Finding a *solution* means finding suitable combinations of causes and effects that meet the constraints.
- $\mathbf{C}$ is the set of constraints. Constraints are split as *known* and *unknown* ($\mathbf{C}_k$, $\mathbf{C}_u$). The former are evaluated during the search by the algorithm (e.g., causes that cannot occur together). The latter encode sets of constraints not necessarily known a priori: because of this, proposed solutions can be just **hypotheses** (like in abductive reasoning) and need to be assessed for their *plausibility*, as they could violate the unknown constraints.
- π, ν. A solution is characterized by a *plausibility* and a *novelty* score, $\pi(\mathbf{x})$ and $\nu(\mathbf{x})$, which are the objective functions ($\pi : \Omega \to \Pi \subseteq \mathbb{R}$ and $\nu : \Omega \to N \subseteq \mathbb{R}$, with Π and N taking values in $[0; 1]$). To assess $\pi(\mathbf{x})$ and $\nu(\mathbf{x})$, a CCOP requires the use of a knowledge base *KB*, i.e., a set of cause-effect combinations (i.e., $\{x_{s_1}, ..., x_{s_j}\}$, $\{x_{t_{j+1}}, ..., x_{t_n}\}$) already observed, representing the experience-based knowledge with respect to which plausibility and novelty of generated solutions are assessed.

3 PATTERNS OF ABDUCTION

Following [13], we distinguish *factual* from *creative abduction*, in turn classified in *analogical* and *hypothetical cause abduction*:

Factual abduction: both the effect and the abduced causes are singular *facts* and abduction is driven by known implicational laws going from causes to effects. The set of possible combinations of causes can be generated by backward-chaining inference. In this form, abduction has been studied in detail in AI [8].

Analogical abduction: it abduces a *partially new* hypothesis by projecting knowledge from previous situations in the domain under analysis. The process involves a conceptual abstraction and a mapping between a *source* context (about which the agent has knowledge) and the *target* context (in which the agent is trying to draw inferences). Both factual and analogical abduction use a form of *experience-based knowledge* – the former about the problem to solve (what we called *KB*), the latter also about an external (source) context from which the analogies are drawn.

[1] A causality relation holds if the values of one – *the cause* - can determine or contribute to the other – *the effect.*

Algorithm 1: Evolutionary Abduction

$P^0_{F/A/H} \leftarrow$ getRandomPop(); ▷ *$P_{F/A/H}$: short for P_F, P_A, P_H*
evaluatePop($P^0_{F/A/H}$); ▷ *Plausibility evaluation of all solutions*
evaluatePopConstraints($P^0_{F/A/H}$); ▷ *Novelty ev. of all solutions*
$S^0_{F/A/H}$, $T^0_{F/A/H} \leftarrow$ getAllSourcesTargets($P^0_{F/A/H}$);
▷ *Get all different source/target values from current population*
while *stopping conditions are not satisfied* **do**
  ▷ *Three sequential loops, cycling on P_F, P_A, P_H; t starts from 1*
  for *i=1 to* $|P^t_{F/A/H}|$ **do**
    $\mathbf{x}_{i,t} \leftarrow$ select_solution($P^t_{F/A/H}$, KB_A);
    $\mathbf{y}_{i,t} \leftarrow$ apply_operator($\mathbf{x}_{i,t}$, $P^t_{F/A/H}$, $S^t_{F/A/H}$, $T^t_{F/A/H}$);
    evaluate($\mathbf{y}_{i,t}$); ▷ *Plausibility evaluation*
    evaluateConstraints($\mathbf{y}_{i,t}$); ▷ *Novelty evaluation*
    $P^{t+1}_{F/A/H} \leftarrow P^{t+1}_{F/A/H} \cup (\mathbf{y}_{i,t})$;
  $Q_{F/A/H} \leftarrow$ nextPopulation($P^t_{F/A/H} \cup P^{t+1}_{F/A/H}$);
  ▷ *Merge population and offspring by non-dominated sorting*
  $t \leftarrow t + 1$; $P^t_{F/A/H} \leftarrow Q_{F/A/H}$;
  $S^t_{F/A/H}$, $T^t_{F/A/H} \leftarrow$getAllSourcesTargets($P^t_{F/A/H}$);
$P \leftarrow P^t_F \cup P^t_A \cup P^t_H$; return $R \leftarrow$ getRankedSolutions(P);

Algorithm 2: factual_operator(x, S, T)

Input : $\mathbf{x}$, the selected solution; S/T, all distinct sources/targets in the current population; η_F, Novelty index; γ_F, Change index
$t \leftarrow$ selectTarget(T);
$\mathbf{x}$' = {$\mathbf{x}$, t}; ▷ *initialize x' with the same sources as x, and target t*
$c \leftarrow$ Rand(1, η_F); ▷ *Number of changes*
for *i=1 to c* **do**
  $a \leftarrow$ Rand(*add, modify, delete*) ▷ *Action to apply*
  if *a=add* **then**
    $s \leftarrow$ selectSource(η_F); ▷ *η_F: Prob to select from S or from KB*
    addSource($\mathbf{x}$', s);
  if *a=modify* **then**
    removeSource($\mathbf{x}$'); $s \leftarrow$ selectSource(η_F); addSource($\mathbf{x}$', s);
  if *a=delete* **then**
    removeSource($\mathbf{x}$');
return $\mathbf{x}$';

Hypothetical cause abduction: this is the most fundamental kind of creative abduction, where we abduce that one or more intercorrelated phenomena are the effect of a hypothetical (unobservable) cause or common cause. One postulates a new unobservable without exploiting analogies. This includes the "pure speculation" process that sometimes lead to find a solution serendipically. Unlike factual and analogical abduction, this abduction does not presuppose any experience-based knowledge, but just knowledge about the phenomenon in explanation,what we called the *ontology*, Ω.

4 EVOLUTIONARY ABDUCTION

To solve a CCOP, EVA keeps an archive of known solutions (*KB*), and a second archive called *analogical KB* (KB_A), used by an *analogical reasoning* operator. Algorithm 1 reports the main steps: *Initialize.* Initially, three sets of solutions (named *sub-populations*) are created by selecting random values for each variable x_i in **s** and in **t** from their domains D_i and that satisfy the known constraints $\mathbf{C}_k$. These represent solutions from the *factual, analogical* and *hypothetical cause* abduction. Plausibility and novelty of these sub-populations are then evaluated (line 2-3).

Apply operators. At every iteration, a *selection operator* and a specific *abduction operator* (i.e.: *factual, analogical* or *hypothetical cause* operator, mimicking the three abduction patterns) are applied to the corresponding sub-population (lines 7-12). *Selection* takes a solution **x** from the sub-population; the abduction operator builds a new solution **x**' starting from **x**. The three sequential loops generate the *offpsring* sub-populations, evolving independently.

Merge. The current and offspring populations are merged by a crowding distance criterion [5] (line 13). All non-dominated fronts $\mathcal{F}_i$ are obtained from the union of current and offspring population by the fast non-dominated sort algorithm. Then, until the population is filled, the crowding distance is calculated in each $\mathcal{F}_i$ and the corresponding solutions are included in the population.

Plausibility and Novelty evaluation

Plausibility. When a solution is proposed by an operator, its plausibility needs to be assessed. Plausibility is the degree to which a hypothesised solution is judged as *realistic*. Typically, our judgement depends on whether we recognize "parts" of the hypothesis in what we already observed. The plausibility score exploits this notion: whenever we detect co-occurrences of (subsets of) causes and effects of the hypothesised solution $\mathbf{x}_j$ in *KB* at least once, we increase our belief about its plausibility. Specifically:

Definition 1 (***k-degree*** **(δ_k)).** The *k-degree* of $\mathbf{x}_j$ ($\delta_k(\mathbf{x}_j)$) is the number of distinct k-tuples of the source variables set **s** (with $k \leq |\mathbf{s}|$) that occurred at least once in *KB* along with the target t.

Then, the plausibility $\pi(\mathbf{x}_j)$ of $\mathbf{x}_j = \{\mathbf{s}, t\}$ with $p = |\mathbf{s}|$ is:

$$\pi(\mathbf{x}_j) = \frac{\sum_{k=1}^{p} \delta_k(\mathbf{x}_j)}{\sum_{k=1}^{p} \binom{p}{k}} = \frac{\sum_{k=1}^{p} \delta_k(\mathbf{x}_j)}{2^p - 1} \quad (2)$$

which is the ratio of all k-tuples of **s**, excluding the 0-tuple, that occurred at least once in *KB* along with t over all the possibilities.[2]

Novelty. Hypothesised solutions need to be plausible but also *different* from those already observed, to avoid convergence towards solutions already present in *KB*. Therefore, their *novelty* with respect to the *KB* is considered, measured as Jaccard similarity:

Definition 2 (***Novelty***). The novelty $\nu(\mathbf{x}_j)$ of a solution $\mathbf{x}_j$ is given by the minimum dissimilarity of $\mathbf{x}_j$ with respect to all solutions $\mathbf{x}_h$ in *KB*, thus: $\nu(\mathbf{x}_j) = 1 - max_h(J(\mathbf{x}_j, \mathbf{x}_h))$, where: $\mathbf{x}_h \in KB$, with $h = 1$ to $|KB|$, $J(\cdot, \cdot)$ is the Jaccard similarity coefficient.

Operators. Selection is always done with the Deb's version of *Binary Tournament* [5]. The other operators are:

Factual abduction. Algorithm 2 takes: a solution **x**, chosen by `select_factual`; all the different (source and target) variables' values that are in the current population (S and T). To build the new solution **x**', first a target t is selected from T. Selection takes two targets randomly, and selects the one with greatest "support" (number of occurrences in *KB*), with ties broken randomly. As for the sources, the same sources of **x** are initially used (line 2). The operator applies three types of changes to the sources: `add`, `modify` or `delete` actions. Two parameters to regulate the extent of changes and the desired novelty are used: a *change index* $\gamma_F > 0$ (integer) and a *novelty index*, $\eta_F \in [0; 1]$ (double). The number of changes c to apply are selected randomly, with $c \in [1; \gamma_F]$. The action (add, modify or delete) is selected randomly with equal chance (line 5). In case of *add* or *modify*, the new source is selected from S or from *KB* (with probability η_F and $1 - \eta_F$, respectively). Sources selection (lines 7, 11, 14) is done by a variable-level binary tournament.

[2] The so-defined $\pi(.)$ is biased toward small solutions, as the denominator is exponential with p. To account for this, $\pi(.)$ is scaled by a factor p/a (if $p \leq a$) or a/p (if $p > a$), where p is the solution size and a is the average size of solutions in the KB.

Algorithm 3: analogical_operator(x, P, S, T)

Input : **x**, the selected solution (from KB_A), P, population; S/T, all distinct sources/targets in the current pop.; η_A, Novelty index

$t \leftarrow$ selectTarget(T);
$[p, v_g, \sigma_{M_g}]$ = extractConstraints();
▹ *Extract #sources (p), #sources per group (v_g), σ_{M_g} per group*
for *i=1 to p* **do**
  $s \leftarrow$ selectSource(η_A); ▹ *η_A: Prob. to select from S or from Ω*
  addSource(**x**', s);
while *(v_g and σ_{M_g} constraints are not satisfied)* **do**
  replaceSource(**x**'); ▹ *Adjust the solution to meet constraints*
return **x**';

Analogical abduction Alg. 3 selects **x** from KB_A. To build a new solution **x**' from **x**, a target t from T is first selected like in factual abduction. Then, it builds the set of sources, coupled with t trying to have the same *structural* features as the sources in **x** (`extractConstraint`). EVA defines three constraints requiring **x**' to have progressively stronger similarities with **x**: the *cardinality constraint* requires that the number of sources of **x**' is the same as **x**; the *group membership constraint* assumes that sources can be grouped in homogeneous subsets, and requires **x**' to have the same number of subsets with the same cardinalities as **x**; the *ordinal constraint*, requires that **x**' has the same number of subsets with the same *maximum k-degree*: $M(\mathbf{q}) = \arg\max_k \delta_k(\mathbf{q})$, with $\mathbf{q} \subseteq \mathbf{s}$.

Hypothetical cause abduction This operator acts exactly as the factual operator (Alg. 2), but considers Ω rather than *KB* in order to possibly select a new, unseen, source. As consequence, these solutions have higher novelty compared to the factual operator but lower plausibility. The indexes in this case will be $\gamma_H > 0$ and $\eta_H \in [0; 1]$, playing the same role of γ_F and η_F in Alg. 2.

5 EVALUATION

Datasets. We used the same datasets used in [12]: the *Primary Tumor dataset* [15] (18 variables, 339 entries). The *ASRS dataset* about avionics accidents [12], [1] (28 variables, 4,470 entries). The *Diabetes dataset* with diabetes information of 70 patients [9] (14 variables, 3,640 entries). The *Nursery dataset*, with data about applications for nursery schools [10] (9 variables, 12,960 entries). In all the cases, one variable is the effect to predict, the others are possible causes. We set: $|KB|$=10% and $|KB_A|$ = 2.5% of total entries.

Baselines. We compare EVA against the four MOEAs used in the CCOP work [12]. They are variants of conventional MOEAs customized to solve a CCOP: NSGA-II [14], OMOPSO [14], SMS-EMOA [2], SPEA2 [16]. For details on customizations see [12].

Table 1: Results: +, ≈, − indicates that EVA is statistically better, equivalent or worse than the compared algorithm.

Dataset	Metric	EVA (*Best*)	EVA (*Worst*)	NSGA-II	OMOPSO	SMS-EMOA	SPEA2
TUMOR	**HV**	**5.18e-01**$_{2.4e-02}$	5.02e-01$_{4.9e-02}$(≈)	1.24e-01$_{2.5e-03}$(+)	3.58e-02$_{6.4e-02}$(+)	9.84e-02$_{2.5e-02}$(+)	1.23e-01$_{2.6e-02}$(+)
	IGD	**1.42e-04**$_{6.3e-04}$	1.46e-04$_{5.0e-04}$(≈)	2.51e-04$_{1.1e-04}$(≈)	2.43e-03$_{5.8e-04}$(+)	8.86e-04$_{8.8e-04}$(+)	2.51e-04$_{7.5e-04}$(≈)
	$\bar{d}$; d^*	**1.43e-01; 1.11e-01**	1.45e-01 (≈); 8.50e-02 (≈)	8.60e-01 (+); 8.33e-01 (+)	8.30e-01 (+); 7.88e-01 (+)	8.21e-01 (+); 7.95e-01 (+)	8.51e-01 (+); 8.17e-01 (+)
ASRS	**HV**	**8.08e-01**$_{3.7e-02}$	7.96e-01$_{2.5e-02}$(≈)	5.68e-01$_{8.0e-02}$(+)	4.58e-01$_{3.2e-02}$(+)	5.17e-01$_{7.6e-02}$(+)	5.29e-01$_{1.0e-01}$(+)
	IGD	**8.02e-04**$_{1.4e-05}$	8.80e-04$_{9.2e-04}$(≈)	3.56e-03$_{7.9e-04}$(+)	4.70e-03$_{7.8e-04}$(+)	5.38e-03$_{3.4e-03}$(+)	3.69e-03$_{1.2e-03}$(+)
	$\bar{d}$; d^*	**3.09e-01; 2.27e-01**	4.18e-01 (+); 2.72e-01 (≈)	9.14e-01 (+); 8.33e-01 (+)	9.05e-01 (+); 8.32e-01 (+)	8.75e-01 (+); 8.47e-01 (+)	9.10e-01 (+); 8.35e-01 (+)
MEDICAL	**HV**	**7.44e-01**$_{5.6e-03}$	7.18e-01$_{4.6e-03}$(≈)	5.69e-01$_{5.0e-03}$(+)	5.20e-01$_{2.9e-02}$(+)	5.52e-01$_{3.3e-02}$(+)	5.55e-01$_{1.5e-02}$(+)
	IGD	**8.02e-05**$_{1.1e-06}$	2.08e-04$_{3.9e-06}$(+)	5.19e-04$_{7.9e-06}$(+)	1.24e-03$_{3.2e-04}$(+)	8.30e-04$_{3.9e-04}$(+)	9.26e-04$_{4.2e-04}$(+)
	$\bar{d}$; d^*	**3.36e-01; 0e+00**	3.45e-01 (≈); 0e+00 (≈)	7.39e-01 (+); 4.43e-01 (+)	7.12e-01 (+); 3.61e-01 (+)	4.19e-01 (+); 3.38e-01 (+)	7.37e-01 (+); 4.32e-01 (+)
NURSERY	**HV**	**8.95e-01**$_{7.2e-03}$	8.94e-01$_{7.8-03}$(≈)	6.25e-01$_{1.5e-02}$(+)	6.20e-01$_{1.5e-02}$(+)	6.25e-01$_{1.3e-02}$(+)	6.25e-01$_{1.2e-02}$(+)
	IGD	**8.05e-05**$_{4.8e-06}$	1.16e-04$_{8.2e-05}$(+)	1.54e-04$_{1.7e-04}$(+)	5.92e-04$_{5.0e-04}$(+)	2.26e-03$_{2.1e-03}$(+)	1.69e-04$_{2.1e-03}$(+)
	$\bar{d}$; d^*	**1.27e-02; 0e+00**	1.27e-02 (≈); 1.46e-01 (+)	7.37e-01 (+); 5.55e-01 (+)	7.30e-01 (+); 5.53e-01 (+)	5.65e-01 (+); 5.53e-01 (+)	7.25e-01 (+); 5.55e-01 (+)

Metrics. We use the *Hypervolume (HV)* [17] (reference point $p = (0,0)$, worst plausibility and novelty); the *Inverted Generational Distance (IGD)* [3] (with the reference front R computed as union of the reference fronts of compared algorithms); a *Distance* metric d, with which the generated solutions $q \in Q$ are compared against solutions in the test set $z \in Z$ to assess how much they are close to *at least one* real occurrence. We use the *minimum Jaccard distance*: $d_{min}(q) = \min_{z \in Z}(d(q,z))$. A small d_{min} tells that the agent has built a solution similar to a real occurrence. We report the average $\bar{d} = avg(d_{min}(q))$ and minimum $d^* = min(d_{min}(q))$ over $q \in Q$.
Setting. To consider EVA in its best and worst configuration, we run a 3×3 grid search on 10 repetitions: we consider 3 configurations of hyperparameters $(< \eta., \gamma. >= (< 0.1, 3 >, < 0.5, 5 >, < 0.9, 7 >)$, representing *Low, Medium* and *High* novelty, and 3 population sizes, $|P| = (15, 30, 60)$. Two configuration are considered for each datasets, producing the best (**B**) and worst (**W**) distances (results in Appendix). The following results are under these settings (B/W).

The setting for the baseline MOEAs is the default one as provided by the used framework (*jMetal*) [7] (Appendix, Sec. 2). The number of evaluations is the same as the CCOP work [12], $K = 20,000$. The generations g depends on the population size $|P|$ (in turn depending on the **B**/**W** configuration), as $K = |P| \times g$. We run 30 repetitions.
Results. Table 1 reports the median and IQR of HV, IGD, $\bar{d}$, and d^*. The best ones are highlighted gray. The statistical comparison is by the Wilcoxon rank sum test ($\alpha = .05$), with Benjamini-Hochberg correction for multiple-comparison protection. As main findings: *i)* EVA's improvement is largely confirmed over all the metrics and datasets, and with a large gap. Besides EVA, NSGA-II gives, in terms of HV and IGD, the second best values in most cases. Looking at the dataset, we notice a low HV in TUMOR (despite the good distances) due to a scarce diversity of produced solutions. *ii)* Considering the distances, all the baselines are significantly worse than EVA; the best second values of $\bar{d}$ are produced by SMS-EMOA, while the best second values of d^* are produced by OMOPSO. Thus, SMS-EMOA had better average performance, while OMOPSO had better smallest-distance solutions (i.e., the close-to-real solutions are fewer but are closer). *iii)* Conventional MOEAs are inadequate to give close-to-real solutions, especially for complex problems (in ASRS, EVA largely outperforms). The gap is significant almost always. *iv)* A final remark about NSGA-II: although very far from EVA, it gives many solutions with good $\bar{d}$, but not in terms of d_{min}; this partly explains the larger HV/IGD than the others, but worse distances.

6 CONCLUSION

This article introduced EVA, a new algorithm for evolutionary computation inspired by abductive inference. EVA is a first step toward a better understanding of how mimicking human reasoning can support in creating algorithms for solving optimization problems.

ACKNOWLEDGEMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 871342

REFERENCES

[1] Last checked: Feb. 2023. NASA Aviation Safety Reporting System. (Last checked: Feb. 2023). https://asrs.arc.nasa.gov/index.html
[2] N. Beume, B. Naujoks, and M. Emmerich. 2007. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *Eur. J. Oper. Res.* 181 (2007), 1653–1669.
[3] P. A. N. Bosman and D. Thierens. 2003. The balance between proximity and diversity in multiobjective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation* 7, 2 (2003), 174–188.
[4] J. A. Crowder, J. Carbone, and S. Friess. 2020. *Artificial Creativity and Self-Evolution: Abductive Reasoning in Artificial Life Forms.* Springer, 65–74.
[5] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. on Evolutionary Computation* 6, 2 (2002), 182–197.
[6] Igor Douven. 2017. Abduction. In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (Ed.). Metaphysics Research Lab, Stanford University.
[7] J. J. Durillo and A. J. Nebro. 2011. jMetal: A Java framework for multi-objective optimization. *Advances in Engineering Software* 42, 10 (2011), 760 – 771.
[8] P. Flach and A. Kakas. 2000. *Abduction and induction.* Dordrecht: Kluwer.
[9] A. Frank and A. Asuncion. 2010. UCI Machine Learning Repository. (2010). http://archive.ics.uci.edu/ml
[10] M. Bohanec M. Olave, V. Rajkovic. 1989. *Expert Systems in Public Administration.* Elsevier Science Publishers (North Holland), Chapter An application for admission in public school systems, 145–160.
[11] Charles S. Peirce. 1931. *The Collected Papers of Charles Sanders Peirce.* Harvard University Press, Cambridge.
[12] Roberto Pietrantuono. 2021. Automated Hypotheses Generation via Combinatorial Causal Optimization. In *2021 IEEE Congress on Evolutionary Computation (CEC).* 399–407.
[13] G. Schurz. 2008. Patterns of Abduction. *Synthese* 164, 2 (2008), 201–234.
[14] M. R. Sierra and C. A. Coello Coello. 2005. Improving PSO-Based Multi-objective Optimization Using Crowding, Mutation and âĹĹ-Dominance. In *Evolutionary Multi-Criterion Optimization*, C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler (Eds.). Springer Berlin Heidelberg, 505–519.
[15] Ljubljana Yugoslavia University Medical Centre, Institute of Oncology. Last checked: Feb. 2023. Primary Tumor Data Set. (Last checked: Feb. 2023). https://archive.ics.uci.edu/ml/datasets/primary+tumor
[16] Eckart Zitzler, Marco Laumanns, and Lothar Thiele. 2002. SPEA2: Improving the Strength Pareto Evolutionary Algorithm. In *EUROGEN 2001.* 95–100.
[17] E. Zitzler and L. Thiele. 1999. Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation* 3, 4 (1999), 257–271.