Supplemental Material for the paper "An Evolutionary Strategy for Automatic Hypotheses Generation inspired by Abductive Reasoning"

Anonymous Author(s)

ABSTRACT

The supplementary material for the paper entitled "An Evolutionary Strategy for Automatic Hypotheses Generation inspired by Abductive Reasoning" includes *textual material* and *artifacts*. Textual material is in the following Sections 1-8. Artifacts includes the source code (and the executable `.jar`) of the proposed algorithm and of all the implemented baselines, the experimental code, the datasets used for the experimentation, and the results reported in the main text and in the following appendix. These are available at: https://anonymous.4open.science/r/EVA

The following textual supplementary material is organized as follows. After the reproducibility statement (Section 1), Sections 2 and 3 report the description of the customized MOEAs we borrowed from [11], followed by the description of the baseline strategies we have implemented to solve the problem by causal structure discovery algorithms. Section 4 reports the results of the tuning of the parameters used in the experimentation. These refer to both the EVA hyperparameters and to the size of the population used in the experimental study. A best and worst case for EVA are derived, then used in the final experimentation reported in the main text. Section 5 reports the results achieved by the three abductive operators of EVA, which together contribute to the overall performance of EVA. ÊSection 6 reports the *distribution* of the distances of the last generation's solutions, namely of the final solutions at the end of the algorithm execution. In particular, it reports the distributions of the average and of the best distance of the final population's solutions and of the *relative distance*, a further metric not included in the main text. Section 7 shows how the solutions' best distances of the populations vary with the number of generations, averaged over the 10 repetitions – this is the same graph as in the main text (Figure 1) but referred to *best* distances rather than the *average* distances. Finally, Section 8 details the ASRS dataset, which, unlike the other datasets, is prepared from scratch starting from the ASRS database.

ACM Reference Format:

Anonymous Author(s). 2023. Supplemental Material for the paper "An Evolutionary Strategy for Automatic Hypotheses Generation inspired by Abductive Reasoning". In *Proceedings of The Genetic and Evolutionary Computation Conference 2023 (GECCO '23).* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

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

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 REPRODUCIBILITY

The material reported in this paper, including:

- The datasets used for the experimentation;
- The source code (and the executable `.jar`) of the proposed algorithm and of the implemented baselines; the experimental code;
- The results reported in the main text as well as in the appendixes;

is available at:
`https://anonymous.4open.science/r/EVA`

Instructions are provided in the repository to *reproduce* the same results of the paper, as well as to *replicate* the study with other datasets. Textual configuration files allow to select the datasets, to set EVA hyperparameters and experimental parameters (e.g., population size, novelty constraint), to set the initial seed (leaving the default and specifying 10 runs will reproduce the same result of the paper), to set the split (knowledge base and test set, leaving the default will will reproduce the same result of the paper). Bash scripts named `run.sh` expedite the process of reproducing the results of the paper, with one script in each dataset's folder for both EVA and for the baselines.

2 MOEA BASELINES AND PARAMETERS SETTING

In the empirical studies, vairants of four multi-objective evolutionary algorithms (MOEA) have been used for comparisons purpose, borrowed from the original work introducing the Combinatorial Causal Optimization Problem (CCOP) [11]. These MOEAs are: csNSGA-II (variant of Non-Sorted Genetic Algorithm II [7]), csOMOPSO (variant of Optimized Multi-objective Particle Swarm Optimization [14]), csSMS-EMOA (variant of $\mathscr{S}$ Metric Selection-Evolutionary Multiobjective Optimisation Algorithms [5]), csSPEA2 (variant of Strength Pareto Evolutionary Algorithm 2 [16]), where the prefix *cs* stands for *causal*. Changes regard the operators, while the algorithm steps are the same as the original algorithms.

CCOP solutions have not a fixed length, as a different number of sources can appear in a solution referring to a target(s). Thus, csNSGA-II, csSMS-EMOA, csSPEA2 adopt a slight variant of the two-point crossover, in which the two crossover points are chosen randomly between 0 and the minimum between the length of the two solutions $\mathbf{x}$ and $\mathbf{y}$ involved, and then the swap operation is performed like in conventional two-point crossover. As for mutation, they adapt a swap mutation operator that replaces, with a given probability, an element of the solution with another.

Changes to csOMOPSO are more substantial. In a OMOPSO algorithm, there is the notion of speed and position of particles (which are the solutions) that change in a continuous range. At each iteration, the algorithm computes, for each particle, *i)* the

Table 1: MOEA parameters setting

	csNSGA-II	*csOMOPSO*	*csSMS-EMOA*	*csSPEA2*
Crossover prob.	0.9	–	0.9	0.9
Crossover index	40	–	20	20
Mutation prob.	(1/n)	(1/n)	(1/n)	(1/n)
Mutation index	20	–	20	20
Perturbation index	–	0.5	–	–
Population size	100	100	100	100
Archive size/offset	–	100	100	100

new position, *ii)* the speed, and applies the *iii)* mutation operator. These form the new solutions to be evaluated. Let us consider U, the domain of interest, with each element i representing an element that can (not necessarily will) be part of a solution; let us as x_i, with $i = 1, \ldots, n = |U|$, the decision variable associated with element i, that can be either a source or a target variable (x_s or x_t). The set of all possible values that a *source* variable can take is denoted as $D_s = \{D_{s_1}, \ldots D_{s_j}\}$; while target variable take values in the respective (target domains): $D_t = \{D_{t_{j+1}}, \ldots D_{t_n}\}$. csOMOPSO splits the continuous $[0, 1]$ interval of values in n equally spaced ranges R_k, and assigns each range R_k to each element $k \in D_s$ or $k \in D_t$ ($k = 1$ to n), so as a potential value of a decision variable is uniquely represented by a range R_k. In this way, each solution $\mathbf{x}$ is a combination of elements represented by a set of continuous values, that correspond to the position in the PSO terminology. The computation of speed and position, as well as the mutation operator, is then applied to such values like in conventional OMOPSO: if a value falls outside its range R_k, than the corresponding source (or target) variable is replaced in the solution, in favour of the variable represented by the new range. If the value exceeds the $[0, 1]$ range, the variable is neglected by the algorithm (i.e., it is "removed" from the solution), while it can be back if the value becomes again included in an R_k range – a solution in a CCOP, as said, can change its size. As for mutation operators: one-third of solutions undergoes the non-uniform mutation, one-third the uniform mutation and one-third are no subject to mutation.

The setting for the described metaheuristics are the default setting as provided by the framework used for experimentation, *jMetal* [8] – they are reported in Table 1. For selection, all the algorithms adopt binary tournament.

3 GRAPH-BASED BASELINE STRATEGIES AND PARAMETERS SETTING

The graph-based (GB) strategies have been implemented as follows. A Causal Structure Discovery (CSD) algorithm is used to learn the causal structure from the knowledge base KB; the output is directed acyclic graph (DAG) with nodes being the variable and arcs being dependency relation between them [10]. This is exploited to generate solutions proportional to cause-effect strength as described hereafter.

The CSD algorithms, namely FGES [12], RFCI [6], and GFCI [9], are all present in the *py_causal* repository [15][1], which exploits the `Tetrad` toolbox [13][2]. The parameters setting to derive the DAG and the corresponding arc weights are in Table 2 – the default parameters are kept, except the number of bootstraps (i.e., number of resampling) raised to 50 to improve the accuracy. The data type is always "discrete". The description of each field can be found at `http://cmu-phil.github.io/tetrad/manual/`:

Table 2: Parameters setting of GB strategies

	FGES	GFCI	RFCI
scoreId	bdeu-score	bdeu-score	–
testId	–	disc-bic-test	bdeu-test
maxDegree/depth	3	3	3
faithfulnessAssumed	True	True	–
numberResampling	50	50	50
resamplingEnsemble	1	1	1
maxPathLength	–	-1	-1
completeRuleSetUsed	–	False	False
addOriginalDataset	True	True	True

As prior knowledge, we specified (by the `priorKnowledge` parameter) that arcs between causes should be forbidden, as we are interested in arcs between causes and effects. The weights between arcs from causes to the effect obtained for the four datasets (values in the repository, `https://anonymous.4open.science/r/EVA`), which represent the probability that a potential *cause* node is causally related to the *effect* node, are used to generate the solution. An example of so-obtained DAG is in Figure 1, wherein *HS* is the effect (Hypoglycemic symptoms) and all the other variables are potential causes such as "More-than-usual meal ingestion" (MTUMI), "Blood Glucose Measurement Decrease" (BGMD). This is obtained by the RFCI algorithm.

Given the graph and their cause-effect weights $W = \{w_i\}$, $i = 1, \ldots, n$ and n being the number of (source) variables, the implemented generator acts as follows: for each instance to generate *i)* includes variable v_i ($i = 1, \ldots, n$) with probability w_i as part of the solution, and then *ii)* selects a value j of the variable v_i, say $h_{i,j}$, proportionally to the estimate of its probability of occurrence $p_{i,j}$ within the KB – obtained as (normalized) relative frequency of that value within KB.

The random strategy just the variables v_i ($i = 1, \ldots, n$) with equal probability of selection, and then the values $h_{i,j}$ of v_i with equal probability of selection.

4 EVA PARAMETERS TUNING

A grid search approach is adopted for parameters tuning. The EVA hyperparameters are the novelty indexes, η_F, η_A and η_H, and the change indexes γ_F and γ_H of the abduction operators. Both regulate the extent to which solutions are required to be diverse (hence novel) with respect to the KB and to the current population: the higher the $\eta.$ values, the higher the probability of selecting new unseen sources, and the higher the $\gamma.$ the higher the number of modifications that are done to build a (factual or hypothetical-cause) solution. The following configurations are considered: $< \eta., \gamma. >= (< 0.1, 3 >, < 0.5, 5 >, < 0.9, 7 >$, representing, respectively, a *Low* novelty degree in the solution, a *Medium* novelty and a *High* novelty.

Additionally, due to its evolutionary nature, EVA exploits the notion of population of solutions, whose *size* can impact the final

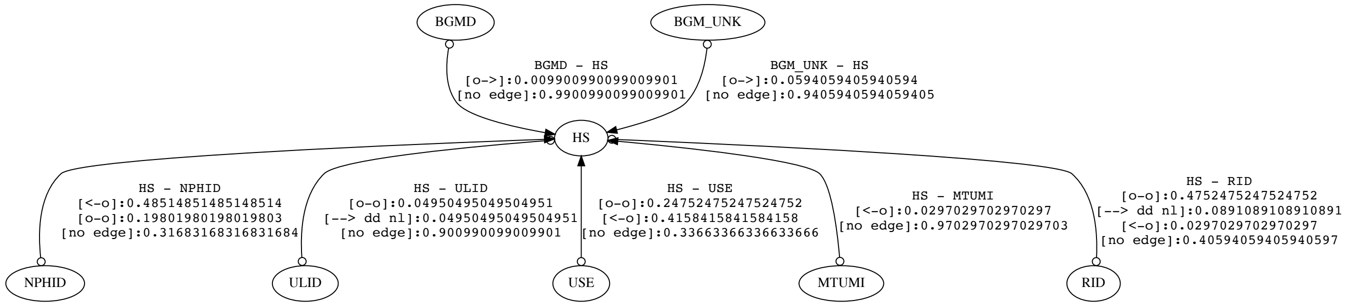


Figure 1: Example of DAG used for the RFCI GB strategy

results. Three values are considered for the population size:$|P| = (15, 30, 60)$.

We ran 10 repetitions for each of the $3 \times 3 = 9$ configurations, each one for 600 evaluations, for the four datasets. Table 3 reports the average distance of the final population’s solution (averaged over the 10 repetitions) from the test set. The **best (B)** and **worst (W)** configurations for EVA are highlighted (green and red, respectively). These two configurations are used to compare EVA with the baselines (over 6,000 evaluations) (cf. with the main paper), considering both the best and the worst case.

5 RESULTS BY EVA OPERATOR

Figure 2 reports the average distance of the final solutions computed by each of the three operators of EVA (i.e.: *Factual, Analogical, Hypothetical-cause*), in every run and experimental scenario (10 runs per scenario) for every dataset.

Two main observations arise: *i)* the *Factual* and *Hypothetical-cause* abduction operators give similar distance values for all scenarios and datasets. These, in fact, have the same structure, the main difference is in the source of knowledge used (the former relies on the KB, while the latter on the ontology Ω); *ii)* the *Analogical* operator works better (i.e., small distances) than the others for small problems, namely when few multiple causes are involved, which is the case of the MEDICAL and NURSERY datasets; in contrast *Factual* and *Hypothetical-cause* outperform the *Analogical* operator for ASRS and TUMOR. The impact of the Best/Worst configuration is negligible, as it does not change the relative results. A higher novelty constraint up to $v_0 = 0.7$ causes the operators’ results to flatten on values above 0.6, as it becomes difficult for all the operators to find close-to-real solutions that are also very different from the KB. The only exception is the case of NURSERY, where the analogical operator still manage to give solutions with distance around 0.5 even with such a strict constraint on the novelty.

Although in one specific problem one operator may provide better solutions, for EVA to work reasonably well with various problems of different size, the suggested strategy is to always exploit the contribution of all the three operators. This also ensures a better diversity of the obtained solutions.

6 DISTRIBUTION OF SOLUTIONS DISTANCE

Figure 3 reports the percentage of solutions of the final generation’s population with average distance less than or equal to a given value – the average over 10 repetitions is reported. For the baseline strategies, since there is no notion of “evolution” and runs (i.e., generations) are independent of each other, we do not consider the final generation, but select the generation with the best population (i.e., having solutions with the best average distance). Results are broken down by novelty constraint and by configuration (best: **B**, worst: **W**).

EVA generates considerably more solutions in the left side of the histogram (i.e., closer to 0) for all the cases. In terms of datasets, the gain is more evident for more complex problems (ASRS, TUMOR), but also for problems with a may instances in the test set (NURSERY), while it becomes less evident for MEDICAL. Again, the Best/Worst configuration makes no relevant difference. With the increase of the novelty constraint the gain of course reduces, as there is less margin for improving over a random or graph-based strategy.

Figure 5 reports the same results but for the *relative* distance. There are many cases in which solutions with relative distance equal to 0 are generated, namely solutions in which the set of causes is entirely contained in the se of causes of a real occurred event (an entry in the test set). For instance, in the MEDICAL dataset, many of the generated solutions (by all the techniques) have relative distance equal to 0[1]. The gain of EVA in terms of relative distance is when the novelty constraint is at 0.1 and 0.4, not at 0.7.

7 BEST SOLUTIONS BY GENERATION

Figures 6 reports the *best* distance of the population’s solutions *vs.* generations (median and IQR over 10). These are the same type of graph as Figure 1 in the main paper, but here the best solution of the population at every generation is considered. The evolution across generations leads to the final results summarized in Table 2 of the main paper. The distances are of course smaller than the average distances of Figure 1 of the main paper. In the case of ASRS, EVA gives distances that still decrease after 6,000 evaluations – it can still improve in that case, while in other cases it converged. When EVA is not visible in the graph (e.g., MEDICAL and NURSERY) it means the distances are 0. Finally, in the case of $v_0 = 0.7$, it often happens that the baselines do not provide solutions for some generations.

[1]Note that the objective is not to generate solutions with small relative distance, in which case would be enough to generate small solutions, e.g., with one single cause. The objective is to generate solutions with small absolute distance; this graph shows how often the so-generated solutions have small relative distance.

Table 3: Average distance (standard deviation) of solutions of the best population – mean over 10 repetitions

		$\lvert P\rvert = 15$	$\lvert P\rvert = 30$	$\lvert P\rvert = 60$
TUMOR	Low	$0.1680_{0.0238}$	$0.1821_{0.0221}$	$0.2118_{0.0155}$
	Medium	$0.1648_{0.0240}$	$0.1755_{0.0154}$	$0.2099_{0.0205}$
	High	$0.1620_{0.0252}$	$0.1652_{0.0335}$	$0.2034_{0.0332}$
ASRS	Low	$0.5407_{0.0259}$	$0.5833_{0.0118}$	$0.6376_{0.0185}$
	Medium	$0.4803_{0.0203}$	$0.5204_{0.0181}$	$0.5961_{0.0191}$
	High	$0.5162_{0.0178}$	$0.4971_{0.0314}$	$0.5560_{0.0196}$
MEDICAL	Low	$0.3584_{0.0209}$	$0.3651_{0.0197}$	$0.3631_{0.0175}$
	Medium	$0.3629_{0.0212}$	$0.3421_{0.0169}$	$0.3677_{0.0134}$
	High	$0.3557_{0.0341}$	$0.3582_{0.0179}$	$0.3615_{0.0108}$
NURSERY	Low	$0.0742_{0.0479}$	$0.0806_{0.0306}$	$0.1292_{0.0193}$
	Medium	$0.0850_{0.0270}$	$0.1101_{0.0309}$	$0.1517_{0.0297}$
	High	$0.1253_{0.0187}$	$0.1336_{0.0300}$	$0.1723_{0.0196}$

8 THE ASRS DATASET

While the TUMOR, MEDICAL and NURSERY datasets were already publicly available and explained, the ASRS dataset is new. Here we briefly describe the source of information from which the dataset is derived.

The Aviation Safety Reporting System (ASRS) database is the world's largest repository of voluntary, confidential safety information provided by aviation personnel, including pilots, controllers, mechanics, flight attendants and dispatchers [4].

It contains more than 1 million of entries reported since 1988. It is a structured database used for data retrieval and analysis, with all the accidents stored in a cause-effect style: the events regarding the aircraft components, the weather conditions, the human personnel involved, the airport, and many other potential causes recorded for each accident as a categorised set of values (i.e., enumerative), along with the resulting accident (also categorised). The main entities are reported in the following:

- **Environment**, with information regarding the flight conditions when accident occurred, visibility, working environment factors such as lighting or temperature.
- **Aircraft**-related elements, e.g., the flight plan, the route, the flight phase, the maintenance status, the mission.
- **Component**, with information about all the components of the aircraft and their status (e.g., design problem, failed, malfunctioning).
- **Person**, reporting the information about the persons involved, such as the flight crew, the air traffic control, or people working in maintenance, information about the human factors that could cause mistakes such as distraction, confusion, stress, etc.
- **Events**, including anomalies such as airspace violation, deviation of altitude, procedural errors, airbone or ground conflict, fire, as well as the event describing the final result, such as the type of accident and its consequences (which correspond to our target variables).

An excerpt of the main information is reported in the Tables 4-8. A glossary of terms is available on the website [3]. For illustrative purpose, a solution looks like follows:

```
Environment.Weather = Fog
Environment.Weather = Windshear
Environment.Weather = Turbulence
Environment.FlighConditions = IMC
Environment.Light = Night
Aircraft.Mission = Cargo/Freight
FlightAircraft.Phase =  Final Approach
Anomaly.Inflight Event = Object encountered
Result.Flight Crew = Landed in
Emergency Condition
Result.Aircraft = Aircraft Damaged
```

This describes an accident in which the pilot, while descending to approach for landing (`Final Approach`) during the night and under bad weather conditions (`IMC` stands for Instrument Meteorological Conditions as opposed to Visual Meteorological Conditions), struck a tree branch (`Object encountered`) and damaged the wing. Hence, he diverted to another airport, landing there in emergency conditions. This type combination is what EVA aims to construct by its operators as described in the main article. The dataset is made publicly available in our repository, `https://anonymous.4open.science/r/EVA`

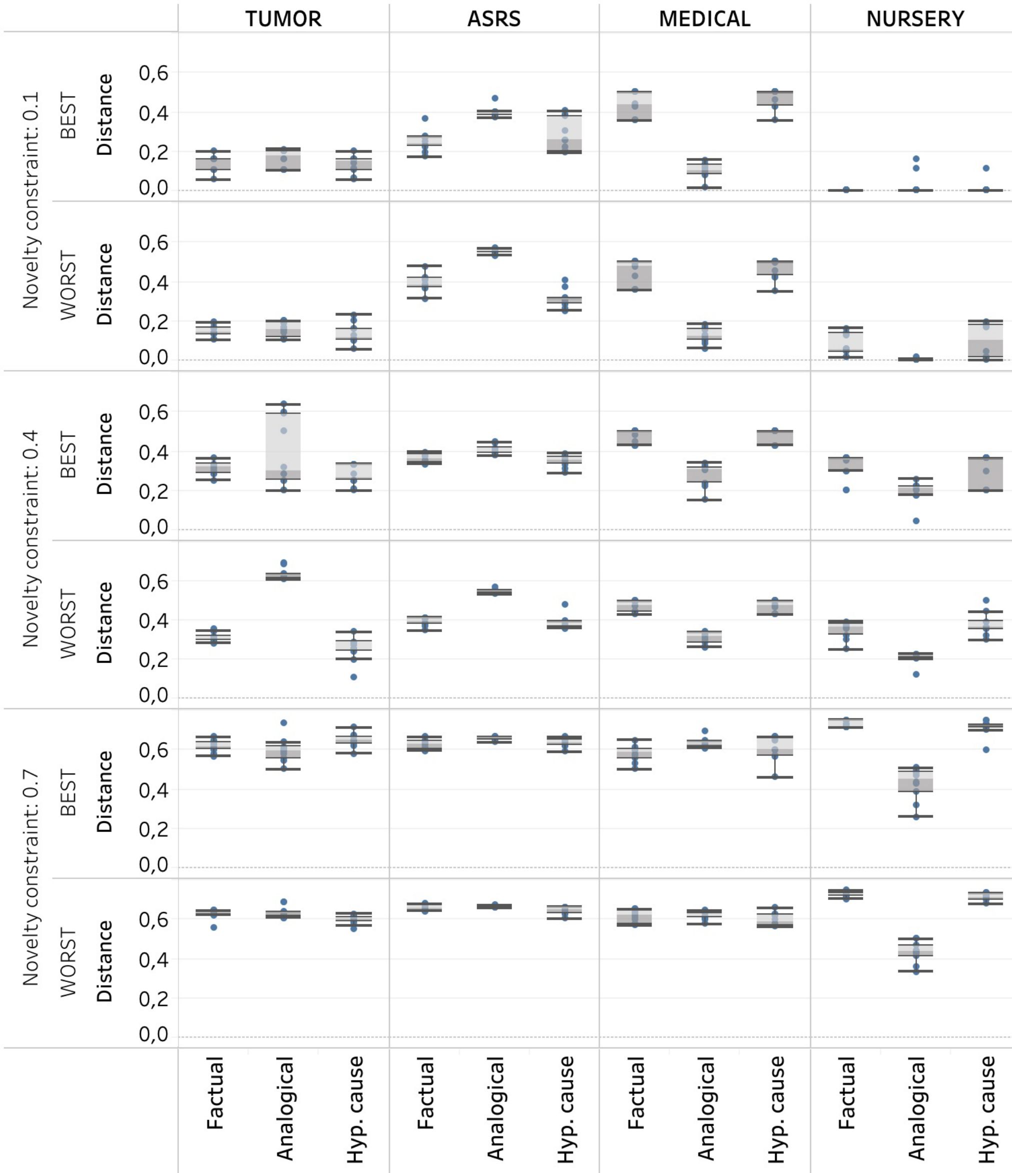


Figure 2: Results by operators

REFERENCES

[1] Last checked: 2022. (Last checked: 2022). https://github.com/bd2kccd/py-causal

[2] Last checked: 2022. (Last checked: 2022). https://www.ccd.pitt.edu/tools/

[3] Last checked: 2022. NASA Aviation Safety Reporting System – Abbreviations. (Last checked: 2022). https://asrs.arc.nasa.gov/docs/dbol/ASRS_Abbreviations.pdf

[4] Last checked: Feb. 2023. NASA Aviation Safety Reporting System. (Last checked: Feb. 2023). https://asrs.arc.nasa.gov/index.html

[5] N. Beume, B. Naujoks, and M. Emmerich. 2007. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *Eur. J. Oper. Res.* 181 (2007), 1653–1669.

[6] Diego Colombo, Marloes H. Maathuis, Markus Kalisch, and Thomas S. Richardson. 2012. Learning high-dimensional directed acyclic graphs with latent and selection variables. *The Annals of Statistics* 40, 1 (2012), 294–321. http://www.jstor.org/stable/41713636

[7] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6, 2 (2002), 182–197.

[8] Juan J. Durillo and Antonio J. Nebro. 2011. jMetal: A Java framework for multi-objective optimization. *Advances in Engineering Software* 42, 10 (2011), 760 – 771. https://doi.org/10.1016/j.advengsoft.2011.05.014

[9] Juan Miguel Ogarrio, Peter Spirtes, and Joe Ramsey. 2016. A Hybrid Causal Search Algorithm for Latent Variable Models. In *Proceedings of the Eighth International Conference on Probabilistic Graphical Models*, Alessandro Antonucci, Giorgio Corani, and Cassio Polpo Campos (Eds.). 368–379.

[10] Judea Pearl. 2009. *Causality: Models, Reasoning and Inference* (2nd ed.). Cambridge University Press, USA.

[11] Roberto Pietrantuono. 2021. Automated Hypotheses Generation via Combinatorial Causal Optimization. In *2021 IEEE Congress on Evolutionary Computation (CEC)*. 399–407. https://doi.org/10.1109/CEC45853.2021.9504816

Table 4: ASRS. *Environment* entity.

Environment				
Flight conditions	Weather Elements/ Visibility	Work Env. Factors	Light	Ceiling
VMC	Cloudy	Poor lighting	Dawn	CLR
IMC	Fog	Glare	Daylight	Single value
Mixed	Hail	Temperature extreme	Dusk	
Marginal	Haze-Smoke	Excessive humidity	Night	
	Icing			
	Rain			
	Snow			
	Thunderstorm			
	Turbolence			
	Windshear			
	Other			

Table 5: ASRS. *Aircraft* entity.

Aircraft						
Flight plan	Flight Phase	Route in use	Navigation in use	Cabin Lighting	Maintenance status & items	Mission
VFR	Taxi	Direct	FMS/FMC	High	Deferred	Aerobatics
IFR	Parked	Oceanic	GPS	Medium	Records	Agricolture
SVFR	Takeoff	VFR Route	INS	Low	complete	Ambulance
DVFR	Initial	Vectors	Localizer/	Off	Released	Banner tow
None	climb	Visual appr.	Gideslop/ILS		for serv.	Ferry
	Climb	None	NDB		Required	Cargo/Freight
	Cruise	Airway	VOR/VORTAC		Scheduled	Passenger
	Descent	STAR			Unscheduled	Photo shoot
	Initial Appr.	SID				Personal
	Final Appr.	Other			*Maintenance items*	Refueling
	Landing				Inspection	Skydiving
	Other				Installation	Tactical
					Repair	Test Flight
					Testing	Traffic watch
					Work cards	Training
						Utility
						Other

[12] Joseph Ramsey. 2015. Scaling up Greedy Equivalence Search for Continuous Variables. *ArXiv* abs/1507.07749 (2015).

[13] Joseph Ramsey, Kun Zhang, Madelyn Glymour, Ruben Sanchez Romero, Biwei Huang, Immé, Ebert-Uphoff, Savini M. Samarasinghe, Elizabeth A. Barnes, and Clark Glymour. 2018. TETRAD - A TOOLBOX FOR CAUSAL DISCOVERY.

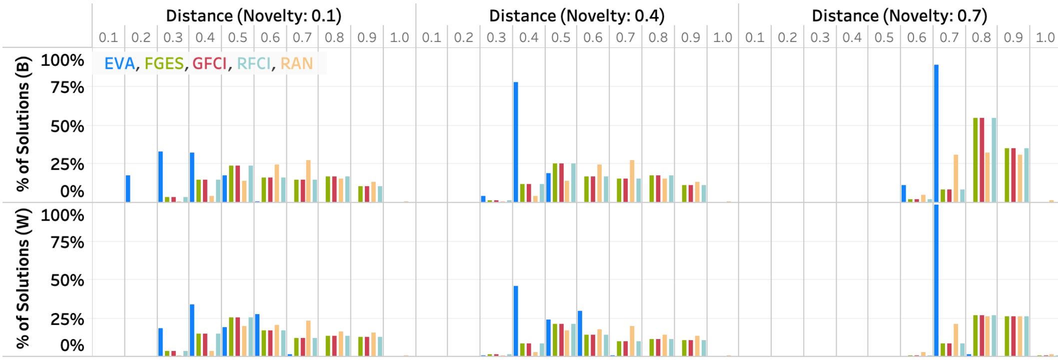
[14] Margarita Reyes Sierra and Carlos A. Coello Coello. 2005. Improving PSO-Based Multi-objective Optimization Using Crowding, Mutation and âĹĹ-Dominance. In *Evolutionary Multi-Criterion Optimization*, Carlos A. Coello Coello, Arturo Hernández Aguirre, and Eckart Zitzler (Eds.). Springer Berlin Heidelberg, 505–519.

[15] Matthew J. Vowels, Necati Cihan Camgoz, and Richard Bowden. 2021. D'ya like DAGs? A Survey on Structure Learning and Causal Discovery. (2021). arXiv:cs.LG/2103.02582

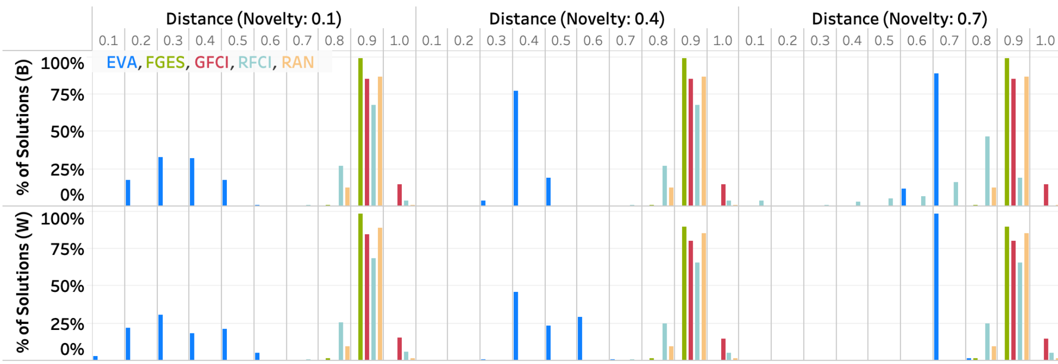
[16] Eckart Zitzler, Marco Laumanns, and Lothar Thiele. 2002. SPEA2: Improving the Strength Pareto Evolutionary Algorithm. In *EUROGEN 2001*. 95–100.

Table 6: ASRS. *Component* entity

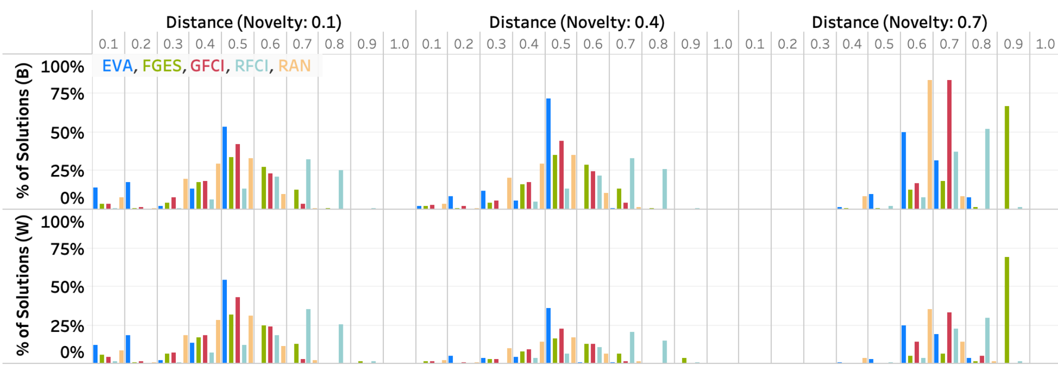
Component		
Component		Problem
Weather Radar	Electrical Wiring & Connectors	Design
DC Battery	Autopilot	Failed
Turbine Engine	Landing Gear	Improperly operated
Indicating and Warning - Landing Gear		Malfunctioning
Nose Gear	Yaw Control	
Flap Vane	Brake System	
Powerplant Fire Extinguishing	Wheels/Tires/Brakes	
Cockpit Window	Aircraft Cooling System	
Turbine Assemb Blade	Landing Gear Indicating System	
Normal Brake System	Tires	
Gear Down Lock	Fuel System	
Engine Control	Fire/Overheat Warning	
Antiskid System	Piston	
Fuselage Skin	Powerplant Fuel Control	
External Power	Flap Control	
Supplemental Landing Gear	FCC (Flight Control Computer)	
Fuselage Panel	*(more than 350)*	
Engine	...	



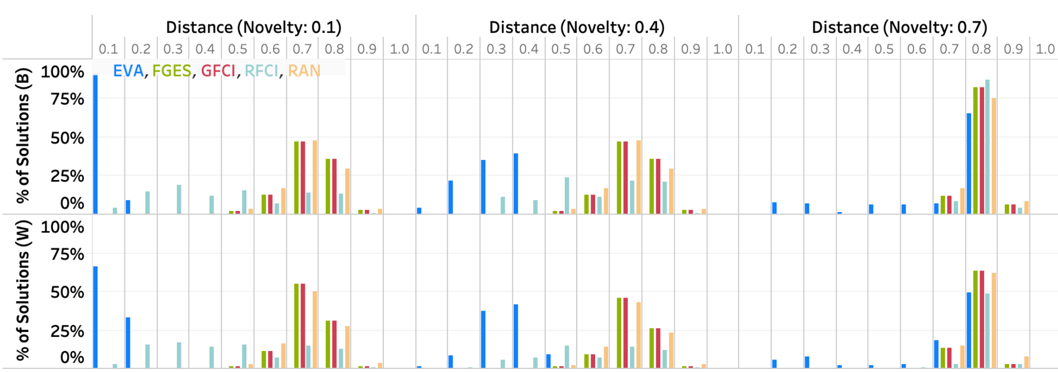
(a) TUMOR problem



(b) ASRS problem

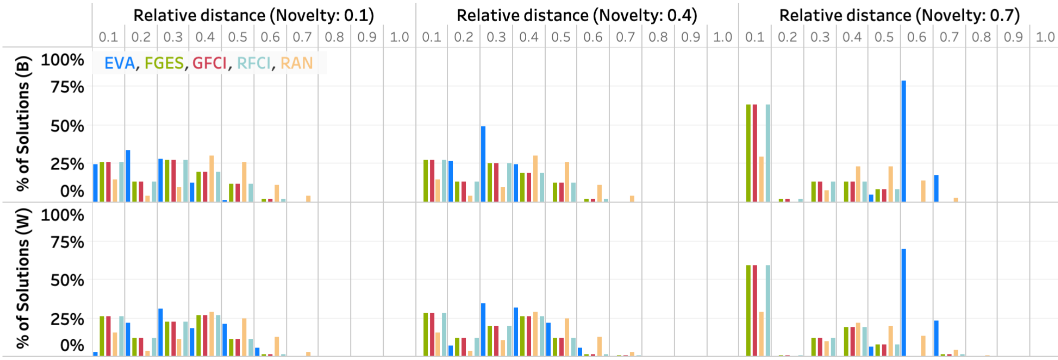


(c) MEDICAL problem

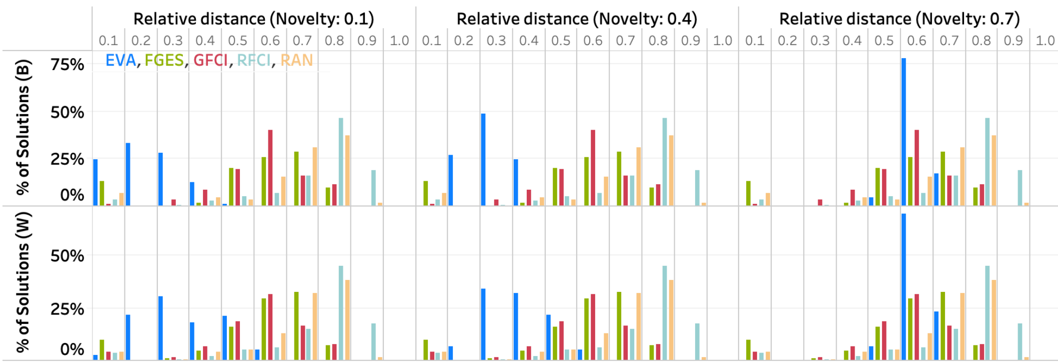


(d) NURSERY problem

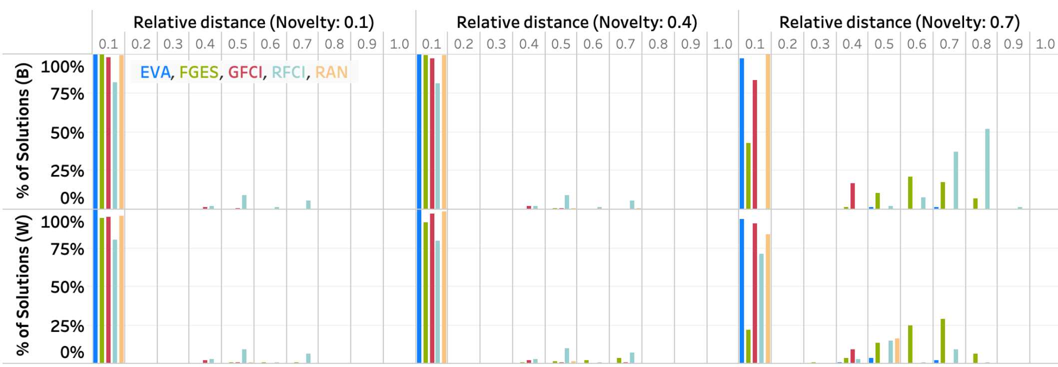
Figure 3: Distribution of solution's distance



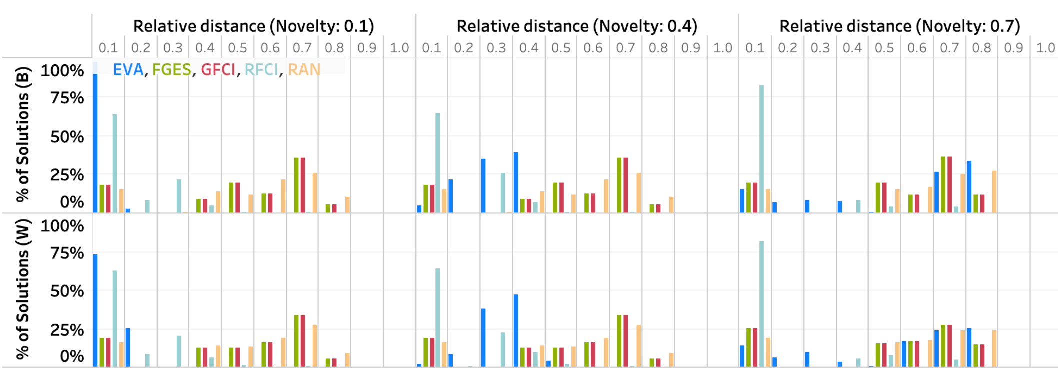
(a) TUMOR problem



(b) ASRS problem

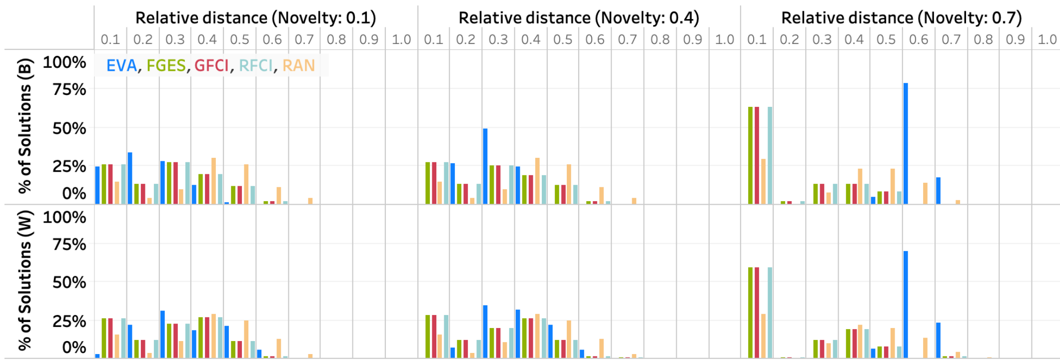


(c) MEDICAL problem

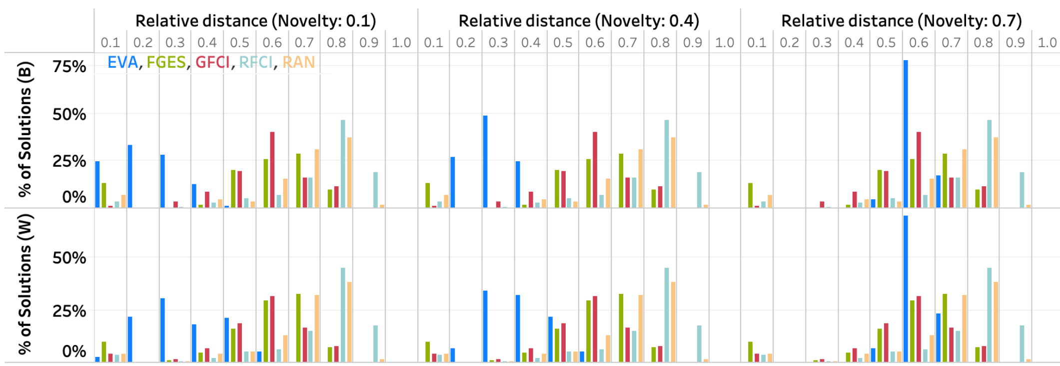


(d) NURSERY problem

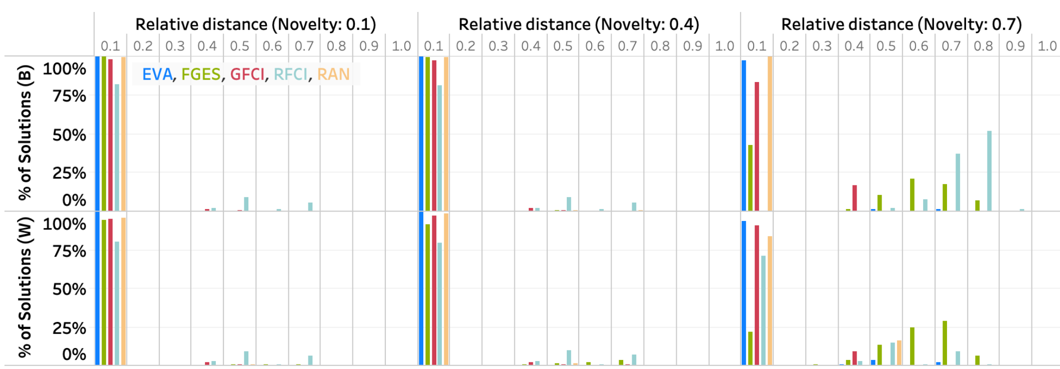
Figure 4: Distribution of solution's best distance



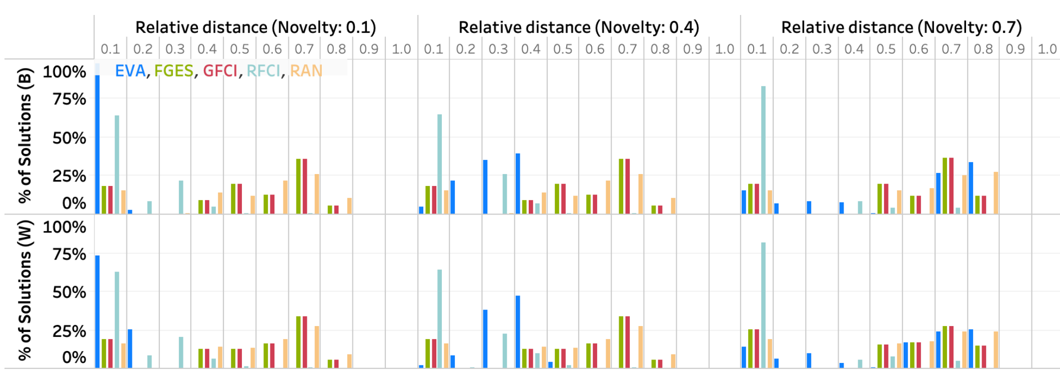
(a) TUMOR problem



(b) ASRS problem

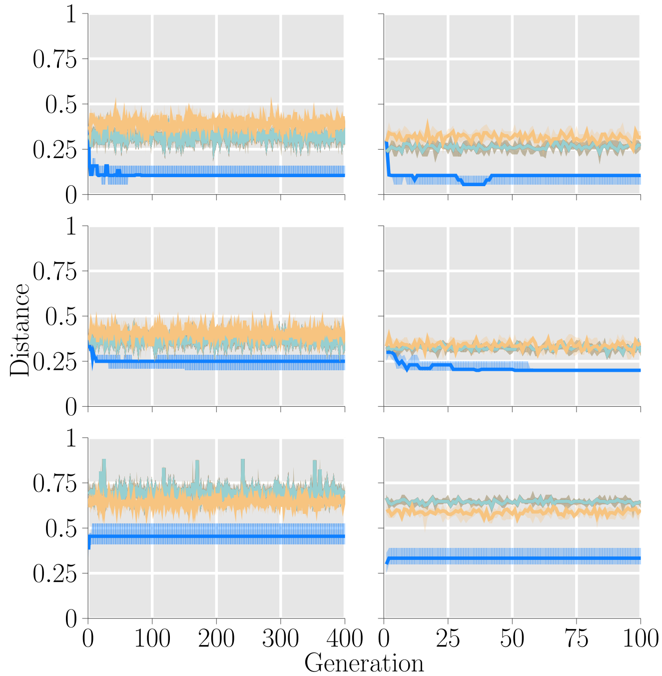


(c) MEDICAL problem

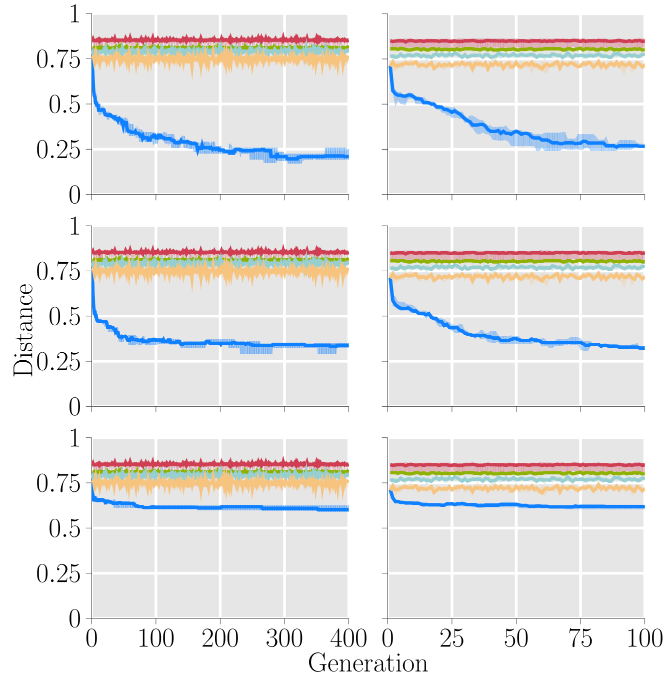


(d) NURSERY problem

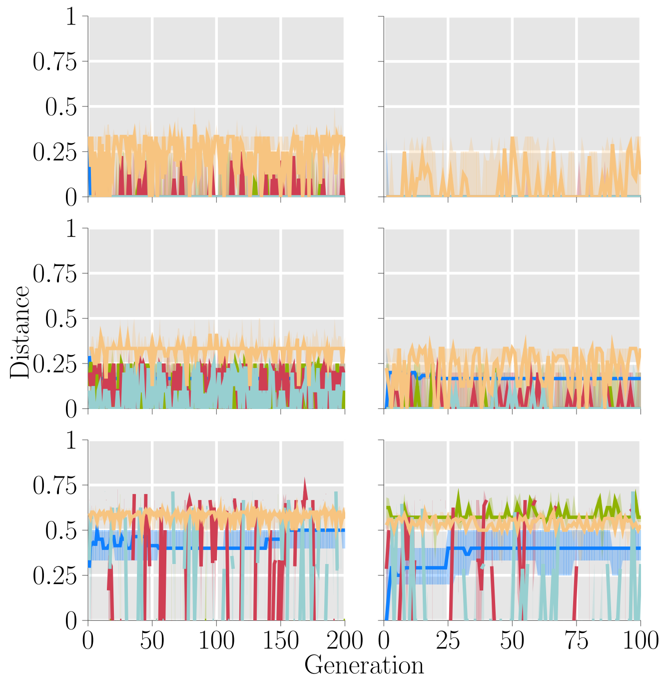
Figure 5: Distribution of solution's relative distance



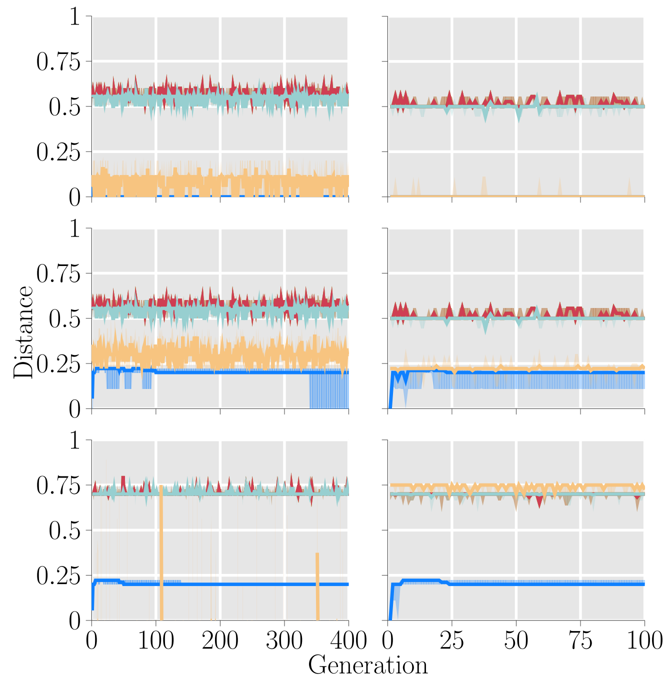
(a) TUMOR problem



(b) ASRS problem



(c) MEDICAL problem



(d) NURSERY problem

EVA FGES GFCI RFCI RAN

Figure 6: *Best distance* by generation.

Table 7: ASRS. *Person* entity

Person			
Function	Qualification	Experience	Human Factors
Flight crew			
Captain	Student	Total	Communication breakdown
Check Pilot	Sport	Last 90 days	Confusion
First Officer	Private		Distraction
Flight Engineer	Commercial		Fatigue
Instructor	Air Transport Pilot		Human-Machine Interaction
Pilot Flying	Flight Instructor		Physiological
Pilot not Flying	Multiengine		Situational Awareness
Relief Pilot	Instrument		Time Pressure
Single Pilot	Flight Engineer		Training/Qualification
Trainee	Rotorcraft		Workload
Other	Lighter-Than-Air		Other
	Sea		
	Glider		*Location in aircraft*
Air Traffic Control			Flight deck
Approach	Fully certified	Radar	Cabin Jumpseat
Coordinator	Developmental	Non-radar	Crew Rest Area
Departure		Military	Dooe Area
Enroute		Supervisory	Galley
Flight data			General Searing Area
Flight service			Lavatory
Ground			Other
Handoff			
Instructor			
Trainee			
Local			
Oceanic			
Supervisor			
Traffic Management			
Other			
Maintenance			
Inspector	Airframe	Avionics	
Instructor	Powerplant	Inspector	
Lead Technician	Appentice	Lead Technician	
Parts/Stores Personnel	Avionics	Repairman	
Quality Assurance	Inspection Authority	Technician	
Technician	Nondestructive Testing		
Trainee	Repairman		
Other			

Table 8: ASRS. *Events* entity

Events		
Anomalies	Assessment Primary or Contributory factor	Results
Aircraft Equipment		*General*
Critical	Aircraft	Declared Emergency
Less severe	Airport	Evacuated
	Airspace structure	Flight Cancelled/Delayed
Airspace Violation	ATC Equip	Maintenance Action
All types	/Nav Facility/Buildings	Physical Injury/Incapacitation
ATC Issues	Chart or Publication	Police/Security Involved
All types	Company Policy	Release Refused/Aircraft not Accepted
Flight Deck/Cabin/Aircraft	Equipment/Tooling	Work Refused
Illness	Env. non-weather related	None
Passenger Electronic Device	Human Factors	*Flight crew*
Passenger Misconduct	Incorrect/Not Instal.	Reoriented
Smoke/Fire/Fumes/Odor	/Unav. Part	Diverted
Other	Logbook Entry	FLC Overrode Automation
Conflict	Manuals	FLC Complied
NMAC	MEL	Executed Go Around/Missed Approach
Airbone conflict	Procedure	Exited Penetrated Airspace
Ground Conflict, critical	Staffing	Inflight Shutdown
Ground Conflict, less severe	Weather	Landed as Precaution
Deviation - Altitude		Overcame Equipment Problem
Crossing Restriction Not Met		Regained Aircraft Control
Excursion from Assigned Altitude		Rejected Takeoff
Overshoot		Requested ATC Assistance/Clarification
Undershoot		Returned to Clearance
Deviation - Speed or Track/Healing		Returned to Departure Airport
All types		Returned to Gate
Deviation - Procedural		Took Evasive Action
Clearance		*Air Traffic Control*
FAR		Provided Assistance
Hazardous Material Violation		Issued Advisory/Alert
Landing without Clearance		Issued New Clearance
Maintenance		Separated Traffic
MEL		*Aircraft*
Published Material/Policy 5205 - Security		Aircraft Damaged
Weight and Balance		Automation Overrode Flight Crew
Other/Unknown		Equipment Problem Dissipated
Ground Excursion/Incursion		
Ramp		
Runaway		
Taxiway		
Ground Event/Encounter		
Aircraft		
FOD		
Gear Up Landing		
Ground Strike Ð Aircraf		
Loss of Aircraft Control		
Object		
Person/Animal/Bird		
Vehicle		
Other		
Inflight Event/Encounter		
CFTT/CFIT		
Fuel Issue		
Loss of Aircraft Control 5215 - Object		
Bird/Animal		
Unstabilized Approach		
VFR in IMC		
Wake Vortex Encounter		
Weather/Turbulence		