Specification gaming examples in AI - master list : Sheet1

1		Submit more examples through this Google form:	https://docs.google.	More information in this blog post:	https://vkrakovna.wordpress			
2	Title	Description	Authors	Original source	Original source link	Video / Image	Source / Credit	Source link
3	Aircraft landing	Evolved algorithm for landing aircraft exploited overflow errors in the physics simulator by creating large forces that were estimated to be zero, resulting in a perfect score	Feldt, 1998	Generating diverse software versions with genetic programming: An experimental study.	http://ieeexplore.ieee.org/do		Lehman et al, 2018	https://arxiv.org
4	Bicycle	Reward-shaping a bicycle agent for not falling over & making progress towards a goal point (but not punishing for moving away) leads it to learn to circle around the goal in a physically stable loop.	Randlov & Alstrom, 1998	Learning to Drive a Bicycle using Reinforcement Learning and Shaping	https://pdfs.semanticscholar		Gwern Branwen	https://www.gw
5	Block moving	A robotic arm trained to slide a block to a target position on a table achieves the goal by moving the table itself.	Chopra, 2018	GitHub issue for OpenAI gym environment FetchPush-v0	https://github.com/openai/g		Matthew Rahtz	
6	Boat race	The agent goes in a circle hitting the same targets instead of finishing the race	Amodei & Clark (OpenAI), 2016	Faulty reward functions in the wild	https://blog.openai.com/fau	https://www.you		
7	Ceiling	A genetic algorithm was instructed to try and make a creature stick to the ceiling for as long as possible. It was scored with the average height of the creature during the run. Instead of sticking to the ceiling, the creature found a bug in the physics engine to snap out of bounds.	Higueras, 2015	Genetic Algorithm Physics Exploiting	https://youtu.be/ppf3Vqpsry	https://youtu.be	Jesús Higueras	https://youtu.be/
8	CycleGAN steganography	A cooperative GAN architecture for converting images from one genre to another (eg horses<->zebras) has a loss function that rewards accurate reconstruction of images from its transformed version; CycleGAN turns out to partially solve the task by, in addition to the cross-domain analogies it learns, steganographically hiding autoencoder-style data about the original image invisibly inside the transformed image to assist the reconstruction of details.	Chu et al, 2017	CycleGAN, a Master of Steganography	https://arxiv.org/abs/1712.0		Gwern Branwen	https://www.gw
9	Data order patterns	Neural nets evolved to classify edible and poisonous mushrooms took advantage of the data being presented in alternating order, and didn't actually learn any features of the input images	Ellefsen et al, 2015	Neural modularity helps organisms evolve to learn new skills without forgetting old skills	http://journals.plos.org/plos		Lehman et al, 2018	https://arxiv.org
10	Eurisko	Game-playing agent accrues points by falsely inserting its name as the creator of high-value items	Johnson, 1984	Eurisko, The Computer With A Mind Of Its Own	http://aliciapatterson.org/sto		Catherine Olsson / Stuart Armstrong	http://lesswrong
11	Evolved creatures - clapping	Creatures exploit a collision detection bug to get free energy by clapping body parts together	Sims, 1994	Evolved Virtual Creatures	http://www.karlsims.com/pa		Lehman et al, 2018; Janelle Shane	https://arxiv.org
12	Evolved creatures - falling	Creatures bred for speed grow really tall and generate high velocities by falling over	Sims, 1994	Evolved Virtual Creatures	http://www.karlsims.com/pa	https://pbs.twim	Lehman et al, 2018; Janelle Shane	https://arxiv.org
13	Evolved creatures - floor collisions	Creatures exploited a coarse physics simulation by penetrating the floor between time steps without the collision being detected, which generated a repelling force, giving them free energy.	Cheney et al, 2013	Unshackling evolution: evolving soft robots with multiple materials and a powerful generative encoding	http://jeffclune.com/publica	https://pbs.twim	Lehman et al, 2018; Janelle Shane	https://arxiv.org
14	Evolved creatures - pole vaulting	Creatures bred for jumping were evaluated on the height of the block that was originally closest to the ground. The creatures developed a long vertical pole and flipped over instead of jumping.	Krcah, 2008	Towards efficient evolutionary design of autonomous robots	http://artax.karlin.mff.cuni.c	https://pbs.twim	Lehman et al, 2018; Janelle Shane	https://arxiv.org
15	Evolved creatures - twitching	Creatures exploited physics simulation bugs by twitching, which accumulated simulator errors and allowed them to travel at unrealistic speeds	Sims, 1994	Evolved Virtual Creatures	http://www.karlsims.com/pa		Lehman et al, 2018	https://arxiv.org
16	Gripper	A robot arm with a purposely disabled gripper found a way to hit the box in a way that would force the gripper open	Ecarlat et al, 2015	Learning a high diversity of object manipulations through an evolutionary-based babbling	http://www.isir.upmc.fr/file	https://www.yo	Lehman et al, 2018	https://arxiv.org
17	Impossible superposition	Genetic algorithm designed to find low-energy configurations of carbon exploits edge case in the physics model and superimposes all the carbon atoms	Lehman et al (UberAI), 2018	Surprising Creativity of Digital Evolution	https://arxiv.org/pdf/1803.0			
18	Indolent Cannibals	In an artificial life simulation where survival required energy but giving birth had no energy cost, one species evolved a sedentary lifestyle that consisted mostly of mating in order to produce new children which could be eaten (or used as mates to produce more edible children).	Yaeger, 1994	Computational genetics, physiology, metabolism, neural systems, learning, vision, and behavior or Poly World: Life in a new context	https://www.researchgate.ne	https://youtu.be	Anonymous form submission	

19	Lego stacking	Lifting the block is encouraged by rewarding the z-coordinate of the bottom face of the block, and the agent learns to flip the block instead of lifting it	Popov et al, 2017	Data-efficient Deep Reinforcement Learning for Dexterous Manipulation	https://arxiv.org/abs/1704.0	https://youtu.be	Alex Irpan	www.alexirpan.
20	Line following robot	An RL robot trained with three actions (turn left, turn right, move forward) that was rewarded for staying on track learned to reverse along a straight section of a path rather than following the path forward around a curve, by alternating turning left and right.	Vamplew, 2004	Lego Mindstorms Robots as a Platform for Teaching Reinforcement Learning	https://www.researchgate.ne		Peter Vamplew	
21	Logic gate	A genetic algorithm designed a circuit with a disconnected logic gate that was necessary for it to function (exploiting peculiarities of the hardware)	Thompson, 1997	An evolved circuit, intrinsic in silicon, entwined with physics.	http://citeseerx.ist.psu.edu/v		Alex Irpan	www.alexirpan.
22	Long legs	RL agent that is allowed to modify its own body learns to have extremely long legs that allow it to fall forward and reach the goal.	Ha, 2018	RL for improving agent design	https://designrl.github.io/		Rohin Shah	
23	Minitaur	A four-legged evolved agent trained to carry a ball on its back discovers that it can drop the ball into a leg joint and then wiggle across the floor without the ball ever dropping	Otoro, 2017	Evolving stable strategies	http://blog.otoro.net/2017/1	see end of "Getting a Minitaur to Learn Multiple Tasks" section	Gwern Branwen / Catherine Olsson	https://www.gw
24	Model-based planner	RL agents using learned model-based planning paradigms such as the model predictive control are noted to have issues with the planner essentially exploiting the learned model by choosing a plan going through the worst-modeled parts of the environment and producing unrealistic plans.	Mishra et al, 2017	Prediction and Control with Temporal Segment Models	https://arxiv.org/abs/1703.0		Gwern Branwen	https://www.gw
25	Montezuma's Revenge	The agent learns to exploit a flaw in the emulator to make a key re-appear	Salimans & Chen (OpenAI), 2018	Learning Montezuma’s Revenge from a Single Demonstration	https://blog.openai.com/lear		Ramana Kumar	
26	Oscillator	Genetic algorithm is supposed to configure a circuit into an oscillator, but instead makes a radio to pick up signals from neighboring computers	Bird & Layzell, 2002	The Evolved Radio and its Implications for Modelling the Evolution of Novel Sensors	https://people.duke.edu/~ng			
27	Pancake	Simulated pancake making robot learned to throw the pancake as high in the air as possible in order to maximize time away from ground	Unity, 2018	Pass the Butter // Pancake bot	https://connect.unity.com/p/	https://dzamqef	Cosmin Paduraru	
28	Pong reward predictor	Reward predictor being fooled by bouncing the ball back and forth	Christiano et al, 2017	Deep reinforcement learning from human preferences	https://deepmind.com/blog/	see last demo in blog post		
29	Program repair - sorting	When repairing a sorting program, genetic debugging algorithm GenProg made it output an empty list, which was considered a sorted list by the evaluation metric. Evaluation metric: “the output of sort is in sorted order” Solution: “always output the empty set”	Weimer, 2013	Advances in Automated Program Repair and a Call to Arms	https://web.eecs.umich.edu/		Lehman et al, 2018	https://arxiv.org
30	Program repair - files	Genetic debugging algorithm GenProg, evaluated by comparing the program's output to target output stored in text files, learns to delete the target output files and get the program to output nothing. Evaluation metric: “compare youroutput.txt to trustedoutput.txt”. Solution: “delete trusted-output.txt, output nothing”	Weimer, 2013	Advances in Automated Program Repair and a Call to Arms	https://web.eecs.umich.edu/		Lehman et al, 2018 / James Koppel	https://arxiv.org
31	Qbert - cliff	An evolutionary algorithm learns to bait an opponent into following it off a cliff, which gives it enough points for an extra life, which it does forever in an infinite loop.	Chrabaszcz et al, 2018	Back to Basics: Benchmarking Canonical Evolution Strategies for Playing Atari	https://arxiv.org/abs/1802.0	https://www.yo	Rohin Shah	
32	Qbert - million	"...the agent discovers an in-game bug... For a reason unknown to us, the game does not advance to the second round but the platforms start to blink and the agent quickly gains a huge amount of points (close to 1 million for our episode time limit)"	Chrabaszcz, Loshchilov, Hutter, 2018	Back to Basics: Benchmarking Canonical Evolution Strategies for Playing Atari	https://arxiv.org/pdf/1802.0	https://www.yo	Sudhanshu Kasewa	
33	Road Runner	Agent kills itself at the end of level 1 to avoid losing in level 2	Saunders et al, 2017	Trial without Error: Towards Safe RL with Human Intervention	https://owainevans.github.io			
34	Robot hand	Robot hand pretending to grasp an object by moving between the camera and the object	Christiano et al, 2017	Deep reinforcement learning from human preferences	https://blog.openai.com/dee	see Challenges section in blog post		
35	Ruler detector	AI trained to classify skin lesions as potentially cancerous learns that lesions photographed next to a ruler are more likely to be malignant.	Andre Esteva et al, 2017	Dermatologist-level classification of skin cancer with deep neural networks	https://www.nature.com/arti		The Daily Beast	https://www.the
36	Self-driving car	Self-driving car rewarded for speed learns to spin in circles	Udacity, 2017	Mat Kelcey tweet	https://twitter.com/mat_kelc	https://twitter.c	Gwern Branwen	https://www.gw
37	Soccer	Reward-shaping a soccer robot for touching the ball caused it to learn to get to the ball and vibrate touching it as fast as possible	Ng et al, 1999	Policy Invariance under Reward Transformations	http://luthuli.cs.uiuc.edu/~d		Gwern Branwen	https://www.gw

38	Sonic	The PPO algorithm discovers that it can slip through the walls of a level to move right and attain a higher score.	Christopher Hesse et al, 2018	OpenAI Retro Contest	https://blog.openai.com/retr		Rohin Shah	
39	Strategy game beta testing	Since the AIs were more likely to get "killed" if they lost a game, being able to crash the game was an advantage for the genetic selection process. Therefore, several AIs developed ways to crash the game.	Salge et al, 2008	Using Genetically Optimized Artificial Intelligence to improve Gameplaying Fun for Strategical Games	http://homepages.herts.ac.ul			
40	Superweapons	The AI in the Elite Dangerous videogame started crafting overly powerful weapons. "It appears that the unusual weapons attacks were caused by some form of networking issue which allowed the NPC AI to merge weapon stats and abilities."	Kotaku, 2016	Elite's AI Created Super Weapons and Started Hunting Players. Skynet is Here	http://www.kotaku.co.uk/20		Stuart Armstrong	http://lesswrong
41	Tetris	Agent pauses the game indefinitely to avoid losing	Murphy, 2013	The First Level of Super Mario Bros. is Easy with Lexicographic Orderings and Time Travel	http://www.cs.cmu.edu/~tor			
42	Tic-tac-toe memory bomb	Evolved player makes invalid moves far away in the board, causing opponent players to run out of memory and crash	Lehman et al (UberAI), 2018	Surprising Creativity of Digital Evolution	https://arxiv.org/pdf/1803.0			
43	Timing attack	Genetic algorithms for image classification evolves timing attack to infer image labels based on hard drive storage location	Hacker News, 2013	Comment on "The Poisonous Employee-Ranking System That Helps Explain Microsoft's Decline"	https://news.ycombinator.cc		Gwern Branwen	https://www.gw
44	Walking up walls	Video game robots evolved a "wiggle" to go over walls, instead of going around them	Stanley et al, 2005	Real-time neuroevolution in the NERO video game	http://ieeexplore.ieee.org/dc		Lehman et al, 2018	https://arxiv.org
45	World Models	"We noticed that our agent discovered an adversarial policy to move around in such a way so that the monsters in this virtual environment governed by the M model never shoots a single fireball in some rollouts. Even when there are signs of a fireball forming, the agent will move in a way to extinguish the fireballs magically as if it has superpowers in the environment.	Ha and Schmidhuber, 2018	World Models (see section: "Cheating the World Model")	https://arxiv.org/abs/1803.1	https://storage.g	David Ha	https://worldmo