

NO DEFINITIVE ANSWER

39

MOST DEFENSES AGAINST ADVERSARIAL EXAMPLES THAT HAVE BEEN PROPOSED SO FAR JUST DO NOT WORK VERY WELL AT ALL, BUT THE ONES THAT DO WORK ARE NOT ADAPTIVE. THIS MEANS IT IS LIKE THEY ARE PLAYING A GAME OF WHACK-A-MOLE: THEY CLOSE SOME VULNERABILITIES, BUT LEAVE OTHERS OPEN.

Ian Goodfellow

NO DEFINITIVE ANSWER