The AI Dilemma:
Growth vs Existential Risk

Difficulty: Hard
Total Points: 14
Author: Zonglun

The link below will direct you to an economic research paper by Stanford professor Dr. Charles Jones. The paper discusses the trade-off between economic growth and the existential crisis brought by Artificial intelligence and provides a simple model. Please carefully read pages 1 to 7 before attempting the following questions.
Primary Link: https://web.stanford.edu/~chadj/existentialrisk.pdf
Alternative Link: https://doi.org/10.1257/aeri.20230570

Question	Score

i. Which of the following is a key **unstated** assumption to the research insights of this paper in relation to finding the optimum value?
   - (a) The value of a year of life in consumption units as a ratio to consumption per person is approximately 6.
   - (b) The utility function used is the Constant Relative Risk Aversion (CRRA) utility.
   - (c) Increasing consumption will always lead to a significant increase in utility.
   - (d) The only cost of adopting Artificial Intelligence is the existential threat it brings.

   MCQ   (3)

ii. Select the option where the optimal number of years A.I. should be adopted (T) is the greatest.
   Assume that the utility function is the Constant Relative Risk Aversion (CRRA) utility, the initial consumption per person is 1, the value of life is 6, and life is always preferred to death.

Probability of Existential Risk (%)	Degree of risk aversion	Annual growth rate of consumption per person (%)

(3)

(a)	1	1	15
(b)	2	1	20
(c)	0.5	2	8
(d)	0.8	1.2	10

MCQ

iii. The utility function is the Constant Relative Risk Aversion (CRRA) utility, the initial consumption per person is 1, the value of life is 6, life is always preferred to death, the probability of existential risk is 1.5%, the degree of risk aversion is 3, and the annual growth rate of consumption per person is 24%. The probability that humanity ends under optimal conditions is E (%)

In this case, what is $\lfloor 100E \rfloor$?

$\lfloor \rfloor$ denotes the floor function—$\lfloor 100E \rfloor$ is the greater integer number less than or equal to $100E$.

Integer (8)

Solution

Solution to (i)

Answer: D

Explanation:

   (a) is an approximate, and stated assumption
   (b) is a stated assumption
   (c) is wrong. In fact, for CRRA utility functions, there is a horizontal asymptote to the utility level as consumption approaches infinity.
   (d) is correct. The whole paper leaves out possible external costs to the environment, etc.

Solution to (ii)

Answer: A

Explanation:

The value of a year of life in consumption units as a ratio to consumption per person $v(c)$ is equal to g/δ, and as $v(c)$ increases, the number of years T increases. As γ increases, T decreases. Therefore, T increases with g and decreases with δ, γ

(a) has the greatest $v(c)=g/\delta$ ratio, and also the lowest γ. It is thus the answer.

Sample Question 3

Solution to (iii)

Answer: 286

Explanation:

$$v(c) = \bar{u}c^2 + \frac{1}{-2}$$
$$v(c) = \bar{u}c^2 - 0.5$$
$$\bar{u} = \frac{v(c) + 0.5}{c^2}$$

Initially, $c = 1, v(c) = 6$,

Hence, $\bar{u} = 6.5$

At optimal T,

$$v(c) = \frac{g}{\delta}$$
$$v(c) = \frac{0.24}{0.015} = 16$$

Hence, sub 16 to $v(c) = 6.5c^2 - 0.5$

$$6.5c^2 = 16.5$$
$$c^2 = 2.53846$$

$c = 1.59326$ (rejected the -ve value, as c always >0)

$c = c_0 e^{gT}$ where $c_0 = 1$

$gT = ln(1.59326)$

$$T = \frac{0.46578}{0.24} = 1.9408$$

$$(1 - \frac{E}{100}) = e^{-0.015*1.9408} = 0.97131$$
$$E = (1 - 0.97131)(100) = 2.8692$$
$$100E = 286.92$$

$$\lfloor 100E \rfloor = 286$$