Economic Implications of Software Minds

Steven Kaas
MIRI Visiting Fellow

Steve Rayhawk, Anna Salamon
Machine Intelligence Research Institute

Peter Salamon
San Diego State University

Abstract

Economic growth has so far come from human minds. The future could bring software minds: AIs designed from scratch, or human brains transferred to computer hardware. Such minds could substitute for humans in a wide range of economic activities—including the research and development that are essential to economic growth. Once minds are software products, they can be copied, accelerated, and improved by economic activity. Our goal in the present manuscript is to explore the implications of a mathematical model of an economy heading toward a technological singularity due to such feedback effects. Specifically, we start from a classic endogenous model of economic growth from Romer (1990), which models the technology level A as representing the number of available designs for different producer durables with additively separable effects on output, and which includes a fixed stock of human capital allocated

Kaas, Steven, Steve Rayhawk, Anna Salamon, and Peter Salamon. 2010.
Economic Implications of Software Minds. The Singularity Institute, San Francisco, CA, Aug. 10.

The Machine Intelligence Research Institute was previously known as the Singularity Institute.

endogenously to either the production of goods or the research of new designs. We modify this model by assuming that beyond some level of technology A_θ, the durables produced can substitute for human capital H as $H = H_\theta + h\sum_{i>A_\theta} x_i$, where h is the marginal rate of substitution of product i for human capital and x_i is the amount of durable i produced. Our findings are as follows. First, the model reaches a vertical asymptote in finite time for all of the growing variables (capital, technology level, consumption, etc.). Second, the asymptote has the nature of a simple pole of the form $A(t) = \frac{a}{exp(bt)-c}$ for constants a, b, and c. Third, an explicit bound is calculated for the time required to reach the vertical asymptote from the time when A_θ is reached. Fourth, the conclusions remain qualitatively unchanged even if Romer's main assumption regarding the growth of technology $\dot{A} = \delta \cdot H_A \cdot A$ is weakened to $\dot{A} = \delta \cdot H_A \cdot A^\gamma$, where H_A is the amount of human capital devoted to research and γ is any positive constant.

1. Introduction

Our goal in the present manuscript is to explore the implications of a mathematical model of an economy capable of manufacturing human-equivalent minds, including the implication of a possible technological singularity. As reviewed in Sandberg (2010), there have been many previous models of this development (Sandberg 2010; Hanson, forthcoming, 1994). Our present effort differs from these previous models by using the growth model of Romer (1990), which is arguably the current standard paradigm for endogenous economic growth.

Romer's endogenous growth model assumes optimal behavior of every facet of an economy. Loosely speaking, this means that the price of anything exactly pays for the value it produces, with the exception that inventors capture only the direct value of their invention, and not the value their invention adds to others' future research. The model predicts exponential growth once the amount of human capital H exceeds a threshold value (see Figure 1). Below this value, no research is performed and growth equals

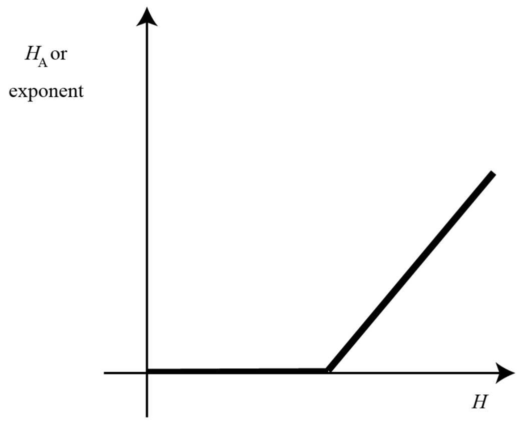

Figure 1: Figure adapted from Romer (1990). The vertical axis shows the exponent for the growth of the economy and the amount of human capital devoted to research, H_A as a function of the total amount of human capital.

zero; above this value, the exponential growth rate increases linearly with H. Thus, it is not surprising that our hypothesis of creating additional human capital H results in super-exponential growth—we move up the line as we produce H. We find that even rather mild hypotheses allowing production of H cause economic output to reach infinity in finite time, provided such production increases H with no upper bound. As argued in Sandberg (2010), such blow-up is not to be taken literally, but rather means that the model predicts a transition to some other regime where the assumptions underlying the model no longer apply.

The main part of Romers model responsible for the blow-up is also present, and also causes a blow-up, in several of the papers reviewed by Sandberg (Sandberg 2010;

Heylighen 2007; Kremer 1993)

$$\frac{dA}{dt} = \delta H_A A \tag{1}$$

where δ is a constant of proportionality, A is the level of technology, and H_A is the amount of human capital devoted to research. For Romer, human capital H is constant. For constant H_A, A grows exponentially, and can be viewed as the driving force behind the exponential growth of the whole economy. This growth dynamic, however, would be completely changed by the possibility of software minds that add to human capital, or even significantly weaker technologies that merely enhance human productivity in such a way as to increase H without bound. Once H_A starts growing as some power v of A, $\frac{dA}{dt}$ grows as A^{1+v} with $v > 0$, and knowledge A, and the total economic output with it, go to infinity in finite time.

2. Reviewing Romer's Model

Technology is modeled by Romer as a sequence of "recipe" indexed by i, $i = 1, ..., A$. The final output Y is

$$Y = H_Y^{\alpha} L^{\beta} \sum_{i=1}^{A} x_i^{1-\alpha-\beta} \tag{2}$$

where H_Y is the human capital used in production, L is labor, and x_i is the number of units of output produced according to recipe i and used in production of durables. Thus, the more recipes that have been produced (via research), the greater the total economic output. The production of consumer goods is modeled only implicitly, by relating production of durables to foregone consumption. Finally, the production of the research sector is given by a slightly rearranged form of equation 1:

$$\Delta A = \delta H_A A \Delta t \tag{3}$$

As stated in Romer (1990):

> An equilibrium for this model will be paths for prices and quantities such that (i) consumers make savings and consumption decisions taking interest rates as given; (ii) holders of human capital decide whether to work in the research sector or the manufacturing sector taking as given the stock of total knowledge A, the price of designs P_A, and the wage rate in the manufacturing sector wA; (iii) final-goods producers choose labor, human capital, and a list of differentiated durables taking prices as given; (iv) each firm that owns a design and manufactures a producer durable maximizes profit taking as given the interest rate and the downward-sloping demand curve it faces, and setting prices to maximize profits; (v) firms contemplating entry into the business of

producing a durable take prices for designs as given; and (vi) the supply of each good is equal to the demand.

Within Romer's model, all durables $i < A$ face the same demand and supply functions, resulting in equal amounts fix of all durables, fetching a price $\overline{p}$ and profit $\overline{\pi} = (\alpha+\beta)\overline{px}$. This profit just pays for the interest on the cost of the invention $\overline{P}_A$ with $\overline{P}_A = \overline{\pi}/r$, where r is the discount rate. Wages for the two sectors are equal, $w_Y = w_A$, and determine the allocation of total human capital $H = H_A + H_Y$, with $H_A = aH - b$, where

$$\alpha = \frac{1}{\delta\Lambda + 1}; \Lambda = \frac{\alpha}{(\alpha+\beta)(1-\alpha-\beta)} \tag{4}$$

with δ equaling the parameter of a Ramsey consumer with utility function

$$U(C) = \frac{C^{1-\delta} - 1}{1-\delta} \tag{5}$$

In particular, note that this has the growth in H_A proportional to growth in H,

$$\Delta H_A = a\Delta H \tag{6}$$

3. Modifying Romer's Model

We extend Romer's model to include the possibility that beyond some level of technology A_θ, the durables produced can substitute for human capital H. Thus, H is no longer constant, but depends on the production levels of those durables that can substitute for human capital. The simplest dependence, and one that calculus assures us must hold for small changes, is a First order Taylor series $H(x) = H_0 + \frac{\partial H}{\partial x}x$. For simplicity, we assume that this local linear approximation holds even for large values of x; thus, our chosen modification of Romer is to take $H = H_\theta + h\sum_{i>A_\theta} x_i$, where $h = h_i = \frac{\partial H_Y}{\partial x_i} > 0$ is the marginal rate of substitution of any product $i > A_\theta$ for human capital and x_i is the amount of durable i produced. As our initial state, we assume an economy following the endogenous exponential growth path solving the classic Romer model. We further assume that the technology reaches a threshold level A_θ, at which we reach designs $i > A_\theta$ that provide recipes for durable goods that increase H by substituting for educated humans in manufacturing and research. We assume also that equation 6 stays valid. While this may not be the optimal allocation for the social planning problem Romer considers, once many recipes $i > A_\theta$ exist, it is enough to make the objective function reach infinity, and thus is as good as any other allocation with this property. The model does not distinguish true software minds from any technologies that extend effective human capital; this may be interpreted as including existing technologies such as Google, Mathematica, etc. insofar as these extend the effective education of researchers.

Once recipes that can be used to increase H have been discovered, all goods are no longer equivalent. Romer's production function implies that smart goods have greater quantities, greater prices and result in greater profits. Thus they are more profitable to develop recipes for and the research sector only develops smart new recipes.

One surprising finding follows from the relationship between P_A and the wage $w = w_A$. In Romer, the entire profit of the research sector becomes wages to the human capital H_A used in this sector. Thus

$$w_A = \frac{P_A \Delta A}{H_A} = \frac{\partial \dot{A}}{\partial H_A} = \delta P_A A \tag{7}$$

Since both P_A and A are growing, Romer's model predicts a growing wage rate. Despite the fact that the amount of human capital tends to infinity, $H \to \infty$, the profitability of each hour of human capital expended toward research also tends to infinity!

A second finding concerns the assumptions required for blow-up. Within Romer's model with constant H, the sustained exponential growth of the economy depends on the assumption that the rate of innovation is proportional to both the amount of human capital and the amount of existing knowledge: $\Delta A \propto H_A A$. Romer notes that this is an assumption rather than a result of his model, and that sustained exponential growth disappears if we do not assume proportionality to A. However, once Romer's model is modified to allow for inventions that substitute for human capital, the assumptions required for unbounded growth become much weaker. Specifically, consider replacing Romer's growth equation with the more general form

$$\frac{dA}{dt} = \delta H_A^v A^\gamma \tag{8}$$

Any time $v + \gamma > 1$, the economic output reaches infinity in finite time. Specifically, this is the case for $v = 1$ and $\gamma > 0$, so it holds even if "low-hanging fruit" for research is depleted much more quickly than in Romer's model. If $v + \gamma = 1$, the production of the research sector shows constant returns to scale and we see exponential growth for the economy, while, if $v + \gamma \leq 1$ (indicating both the exhaustion of low-hanging fruit for research, and limited parallelizability of research), the economic output stays finite and does not blow up. Using $v = \gamma = 1$ as in Romer, and $H_A = u + v(A - A_\theta)$, with constants u and v, we find the solution

$$A(t) = \frac{A_\theta(-u + vA_\theta)}{vA_\theta - ue^{\delta(-u+vA_\theta)t}} \tag{9}$$

which reaches infinity for

$$t = ln(\frac{vA_\theta}{u})\delta^{-1}(-u + vA_\theta)^{-1} \tag{10}$$

The case of interest has $u = H_\theta$ which is indeed constant while $v = h(1-a)\sum_{i>A_\theta} x_i$ is not. In fact v is increasing but such increase merely serves to increase H_A faster and A reaches its vertical asymptote more quickly. It follows that the time t given in equation 10 is only an upper bound and the model actually blows up before t reaches this value.

References

Hanson, Robin. 1994. "If Uploads Come First: The Crack of a Future Dawn." *Extropy* 6 (2). http://hanson.gmu.edu/uploads.html.

———. Forthcoming. "Economic Growth Given Machine Intelligence." *Journal of Artificial Intelligence Research.* Preprint at. http://hanson.gmu.edu/aigrow.pdf.

Heylighen, Francis. 2007. "Accelerating Socio-Technological Evolution: From Ephemeralization and Stigmergy to the Global Brain." In *Globalization as Evolutionary Process: Modeling Global Change,* edited by George Modelski, Tessaleno Devezas, and William R. Thompson, 284–309. Rethinking Globalizations 10. New York: Routledge.

Kremer, Michael. 1993. "Population Growth and Technological Change: One Million B.C. to 1990." *Quarterly Journal of Economics* 108 (3): 681–716. doi:10.2307/2118405.

Romer, Paul M. 1990. "Endogenous Technological Change." In "The Problem of Development": "A Conference of the Institute for the Study of Free Enterprise Systems." Supplement, *Journal of Political Economy* 98 (5, pt. 2): S71–S102.

Sandberg, Anders. 2010. "An Overview of Models of Technological Singularity." Paper presented at the Roadmaps to AGI and the future of AGI workshop, Lugano, Switzerland, March 8th. http://agi-conf.org/2010/wp-content/uploads/2009/06/agi10singmodels2.pdf.