

Re-Thinking the Capabilities of Technology in Economics

Daniel Susskind*

This version: 13 June 2018

Abstract

Over the last 15 years, the ‘task-based’ approach has become dominant in the economic literature that explores the effects of technological change on the labour market. This approach has tended to rely on a particular understanding of the capabilities of technology – known as the ‘ALM hypothesis’. However, this has led this literature to severely underestimate these capabilities. Many tasks that were believed to be out of reach of automation can now be automated. In this paper I set out two distinct explanations for why these capabilities were underestimated – one that is explored in the recent literature and maintains the ALM hypothesis, and a new explanation that challenges it. I propose a new hypothesis about the capabilities of technology that contains the ALM hypothesis as a special case.

Keywords: Technological Change; Automation; Tasks.

JEL: J20; J21; J23.

*Fellow, Balliol College, Oxford University. This is a revised version of Chapter II of my doctoral dissertation, ‘Technology and Employment’. Thanks to David Autor, Brian Bell, Martin Ellison, Frank Levy, Alan Manning, Margaret Stevens, Larry Summers, Richard Susskind, and David Vines for helpful input. Contact: daniel.susskind@balliol.ox.ac.uk.

1 Introduction

Accurately forecasting the future capabilities of new technologies is very difficult. The ‘task-based’ approach, which has become dominant in the economic literature that explores the effects of technological change on the labour market, has often underestimated them. For instance, Autor, Levy, and Murnane (2003) noted that the task of driving a car could not be readily automated, but a type of driverless car appeared two years later;¹ the same paper argued that “legal writing” and “medical diagnosis” could not be reading automated, yet document automation systems are now widespread in most major legal practices and there are a variety of technologies that can diagnose health problems;² Autor and Dorn (2013) noted that order-taking and table-waiting could not be readily automated, but later that year the US restaurants Chili’s and Applebee’s announced they were installing 100,000 tablets to allow customers to order and pay without a human waiter;³ and Autor (2015) noted that the task of identifying a species of bird based on a fleeting glimpse could not be readily automated, but later that year an app was released to do that as well.⁴

These examples suggest that this literature’s understanding of how systems and machines operate and the capabilities that this implies – known as the ‘Autor-Levy-Murnane (‘ALM’) hypothesis’ – is mistaken. There is a growing recognition that this is the case. But what is not yet clear is *why* the ALM hypothesis is mistaken. Explaining why is the main contribution of this paper. First I set out the current explanation in the literature. This revises, but ultimately maintains, the ALM hypothesis. Then I set out a new explanation. This is a fundamental challenge to the ALM hypothesis. It explains why, even in its revised form, the ALM hypothesis is no longer an appropriate way to think about the capabilities of new technologies. In closing, I propose a new hypothesis, based on this second explanation, that nests the ALM hypothesis as a special case. This has implications

¹The Society of Automotive Engineers defines five levels of vehicle ‘autonomy’. These early cars were at a low level. Since 2005, further progress has been made. I thank Frank Levy for this point.

²Automated medical diagnostics is a thriving field. See, for example, Esteva et al. (2017).

³See Pudzer (2016).

⁴See <http://merlin.allaboutbirds.org/photo-id/>.

well beyond the formal economics literature. The ALM hypothesis, despite the limitations set out in this paper, still remains one of the most popular frameworks for thinking about the future of work in expert commentary and among policymakers (for instance, World Bank 2016, ILO 2016, The White House 2016, and IMF 2017).

2 The ALM Hypothesis

The ‘ALM hypothesis’ was introduced in Autor, Levy, and Murnane (2003). Initially, it was developed to explain the capabilities of “computer capital”. But it is now used more generally to think about technology in general. Based on “an intuitive set of observations” from economists and others, the ALM hypothesis is that:

“(1) that computer capital substitutes for workers in carrying out a limited and well-defined set of cognitive and manual activities, those that can be accomplished by following explicit rules (what we term “routine tasks”); and (2) that computer capital complements workers in carrying out problem-solving and complex communication activities (“nonroutine tasks”).”
(pg. 1280)

From this, two questions follow. First, what is the basis for the ‘routine’ vs. ‘non-routine’ distinction? Secondly, what is the basis for the assumption that systems and machines can only perform ‘routine’ tasks?

2.1 ‘Routine’ vs. ‘Non-Routine’ Tasks

In Autor, Levy, and Murnane (2003), the distinction between ‘routine’ and ‘non-routine’ tasks is based on the work of Michael Polanyi. The authors argue that ‘non-routine’ tasks are “tasks fitting Polanyi’s description”, where Polanyi is quoted as follows:

“We can know more than we can tell [p. 4] . . . The skill of a driver cannot be replaced by a thorough schooling in the theory of the motorcar; the knowledge I have of my own body differs altogether from the knowledge of its physiology; and the rules of rhyming and prosody do not tell me what a poem told me, without any knowledge of its rules [p. 20].” [Emphasis added]
(pg. 1283)

Put simply, ‘non-routine’ tasks are those that rely on what Polanyi called ‘tacit’ knowledge – knowledge that people struggle to articulate when called upon to do so.⁵ For example, a doctor may struggle to articulate the rules that allow her to make an insightful diagnosis. But this is still ‘knowledge’ – it is just not ‘explicit’ knowledge that can readily be articulated, but is ‘tacit’. Autor (2014), a more recent re-statement of the task-based approach, again defines ‘non-routine’ tasks with reference to Polanyi’s concept of ‘tacit’ knowledge, arguing that ‘non-routine’ tasks are those we only “tacitly understand how to perform”, that involve processes we “do not explicitly understand”.

2.2 Only ‘Routine’ Tasks Can be Automated

The second part of the ALM hypothesis is the claim that systems and machines can only perform ‘routine’ tasks. This follows from a particular understanding of how systems and machines must operate. In Autor, Levy, and Murnane (2003) this is that a system or machine can only perform a task if it can “follow explicit programmed rules”. This implies that a system or machine requires a task to be “exhaustively specified with programmed instructions”. The same point is made in Autor (2014) but with an additional claim:

“For a computer to accomplish a task, a programmer must first fully understand the sequence of steps required to perform that task, and then must write a program that, in effect, causes the machine to precisely simulate these steps.” (pg. 6)

This claim is not simply that a system or machine must be set an explicit set of programmed rules; it is stronger – that those explicit rules must originate with, and exactly reflect, the thinking process of a human being. It is from this claim that the assumption that computer capital can only perform ‘routine’ tasks follows. Autor (2015) puts this formally:

“But the scope for this kind of substitution is bounded because there are many tasks that people understand tacitly and accomplish effortlessly but for which neither computer programmers nor anyone else can enunciate the explicit “rules” or procedures.” (pg. 11)

⁵In Autor, Levy, and Murnane (2002) there is an informal discussion of this.

And Autor (2014) more prosaically:

“At a practical level, Polanyi’s paradox means that many familiar tasks, ranging from the quotidian to the sublime, cannot currently be computerized because we don’t know the rules.” (pg. 8)

This then explains why, under the ALM hypothesis, ‘non-routine’ tasks cannot be automated – systems and machines have to follow instructions that reflect the rules that human beings follow, but human beings cannot articulate their thought processes for these tasks, so no instructions can be written.

2.3 The Automation of ‘Non-Routine’ Tasks

The ALM hypothesis has a clear implication – systems and machines cannot perform ‘non-routine’ tasks. Yet in practice this no longer holds. The examples before show this. And the recent literature has recognised that ‘non-routine’ tasks can now be automated as well. For instance, the models in Acemoglu and Autor (2011) and Acemoglu and Restrepo (2017; 2018) allow for systems and machines to displace labour in a wider range of tasks than simply ‘routine’ ones.

However, the current explanation for *why* ‘non-routine’ tasks can now be automated still maintains the ALM hypothesis. At present, the prevailing argument is that many ‘non-routine’ task are, in fact, more ‘routine’ than we realised. Advances in processing power, data storage capabilities, and algorithm design, the argument goes, mean that new technologies are able to make ‘explicit’ more of the ‘tacit’ knowledge that human beings draw upon. And that is why ‘non-routine’ tasks can now be automated – human beings are now able to uncover the hitherto hidden tacit rules that they follow in performing a task. This argument is made, for instance, in Autor (2014):

“Contemporary computer science seeks to overcome Polanyi’s paradox [the inability to articulate rules for tasks that require ‘tacit’ knowledge] by building machines that learn from human examples, *thus inferring the rules that we tacitly apply but do not explicitly understand.*” (pg. 2) [Emphasis added]

in Autor (2015):

“[R]ather than teach machines rules that we do not understand, engineers develop *machines that attempt to infer tacit rules* from context, abundant data, and applied statistics.” (pg. 23) [Emphasis added]

and in Remus and Levy (2016) with respect to predicting the outcome of a legal dispute:

“Because the mental protocol [of a judge issuing a decision] is tacit – and not easily articulated – the judge may not experience his decisions as routine, but the *machine learning model makes the tacit protocol explicit* as a mathematical combination of characteristics taken from the case, which can then be used to predict future judicial decisions.” (pg. 13) [Emphasis added]

3 A New Approach

3.1 A New Explanation

However, this current explanation is mistaken. It argues that ‘non-routine’ tasks can now be automated because new technologies are able to uncover more of the tacit rules that human beings follow in performing those tasks. Yet this neglects a critical fact – that many new technologies are performing tasks by deriving and following rules which, on inspection, *do not resemble* the rules that human beings follow at all, tacit or otherwise. This is a new and important conceptual point.

To see this, take the task of making a medical diagnosis. According to the original ALM hypothesis, this task could not be automated because a human doctor, if asked, would struggle to articulate how she performs the task. Yet researchers at Stanford, for instance, have developed a system can predict as accurately as dermatologists whether a picture of a skin discolouration is cancerous (Esteva et al. 2017). According to the revised ALM hypothesis, the reason that this task can now be automated is that this system is able to uncover the hidden rules that doctors follow. Yet this is not right. This system works by running a pattern-recognition algorithm through a database of 129,450 past cases, searching for similarities between

those cases and the particular discolouration in question. There is no reason to think the rules that it derives from that exercise must necessarily resemble those that human beings follow at all.

Or take AlphaGo, a system developed by DeepMind, at Google, to play the Chinese board-game Go. In March 2016, it beat the best human player, Lee Sedol, in a five-game series. Yet it did so not by trying to uncover the tacit rules that human beings play. Instead it derived and followed entirely distinct rules. (The original AlphaGo drew on the data of 30 million past moves by expert human Go players; AlphaGo Zero, the latest version of the system relied on data from three days of self-play alone.) For instance, thousands of years of human tradition had created a rule in Go that stipulated – ‘do not place a stone on the fifth line in from the edge of the board’. This was widely held by human experts to be an awful move. Yet on the 37th move in the 2nd game this is what the system did – and won the game. That successful move has subsequently overturned the rules that human beings follow when playing Go.

These cases are particular examples of a set of techniques known commonly and variously as ‘deep learning’, ‘machine learning’, ‘reinforcement learning’, ‘neural networks’. These systems, rather than follow human rules imposed top-down by programmers from above, instead use processing-power, data storage capabilities, and improvements in algorithm design to derive or ‘learn’ new rules from the bottom-up (for more cases, see Susskind 2016). And critically, these new rules do not necessarily resemble those that human beings follow. Yet the ALM hypothesis, both in its original and revised form, mistakenly assumes exactly this – that the only way to automate a task is to understand, articulate, and replicate the rules that a human being follow in performing that task. The only purpose of the ‘routine’ vs. ‘non-routine’ distinction is to distinguish between those tasks for which human beings can articulate those rules and those they cannot. And the basis for the claim that new technologies can only perform ‘routine’ tasks is that those are the only for which we can write instructions for a machine to follow, since those are the only tasks for which human beings can explain how they perform them. But when new technologies are able to derive different rules to human beings, the ALM hypothesis no longer holds. The inability of human beings to articulate their thinking processes

is no longer such a tight constraint on automation.

More fundamentally, the problem is that the ALM hypothesis draws on very traditional reasoning about ‘artificial intelligence’ (AI). It relies on what is known as the first generation of AI research, conducted from the 1950s to 1980s.⁶ Called ‘expert systems’ or ‘knowledge-based systems’, these required a human ‘domain specialist’ who was capable of articulating how they performed a particular task and then representing this in a set of rules for a machine to follow. Winston (1977), Hayes-Roth et al. (1983) and Russell et al. (1995) provide an overview of these approaches.⁷ But the second generation of AI research, relying on the techniques set out before, no longer requires an articulation of those human rules. These systems and machines derive new rules to follow.

The fact that these new technologies do not necessarily follow rules that resemble the ones that human beings follow has created new problems for AI researchers. In the 1980s, one characteristic of AI research was that systems and machines should be ‘transparent’, able to explain how they reached a particular decision. At the time, this aspiration was relatively easy to achieve – given the rules these early systems followed did reflect the rules that human beings followed, it was straightforward to follow their reasoning. However, today’s systems and machines are far more ‘opaque’ – given they derive and follow rules which do not necessarily reflect the ones that human beings follow, understanding why a system reaches a particular decision is far more difficult. This has led to a new research programme in computer science, focused on developing techniques that allow new systems and machines to ‘explain themselves’ (for instance, the ‘Explainable Artificial Intelligence’ program at DARPA in the US). In the EU, in response, Article 15 of the General Data Protection Regulation (GDPR) has made “meaningful information about the logic” of “automated decision-making” a legal right (see Selbst and Powles 2017).

This analysis does not imply at all that Polanyi’s distinction between ‘tacit’ and ‘explicit’ knowledge is wrong. But it does imply that Polanyi’s distinction is the wrong constraint when thinking about the capabilities of

⁶See Susskind and Susskind (2015).

⁷See also Weizenbaum (1976) and Dreyfus (1976; 1992). These first generation AI researchers also thought a Polanyi-type constraint, based on the inability to express tacit knowledge, was important.

systems and machines that operate in very different ways to human beings. Only when the way in which a system or machine performs a task must exactly reflect the way in which a human being performs that same task does a Polanyi-type constraint bind.

3.2 A New Hypothesis

These two explanations have very different implications for the rules that systems and machines follow in performing ‘non-routine’ tasks. The current explanation argues that these systems and machines are following the ‘tacit’ rules of human beings, which have been made ‘explicit’ – these ‘non-routine’ tasks were ‘routine’ after all. The new explanation I have set out suggests that these new technologies are deriving and following rules that do not reflect the rules that human beings follow – the ‘non-routine’ and ‘routine’ distinction no longer matters.

Given this new explanation, it follows that in thinking about the capabilities of technology the appropriate criteria is not whether a task is ‘routine’ or ‘non-routine’ from the standpoint of a human being, as under the ALM hypothesis, but whether it has features that make it more or less *routinisable* from the standpoint of a system or machine. If a task is routinisable, a routine can be composed that allows a system or machine to perform it – but that routine may not necessarily reflect the way in which a human being performs the task, as the ALM hypothesis assumes. Looked at in this way, the problem with the ALM hypothesis is that it is excessively anthropocentric – it focuses on the rules that *human beings* follow in performing a task, and whether *human beings* find it hard to articulate those rules. The new concept of ‘routinisability’ does neither of these – it recognises that the set of all possible rules a machine could follow is larger than the particular rules that human beings actually do follow. A new general hypothesis follows.

HYPOTHESIS: (1) *that systems and machines substitute for workers in carrying out a set of tasks and activities that are ‘routinisable’; and (2) that systems and machines complement workers in carrying out a set of tasks that are ‘unroutinisable’. The set of tasks and activities that are*

routinisable changes over time.

In the special case where the rules that a system or machine follows must resemble those that human beings follow, the ‘routineness’ and the ‘routinisable’ of a task coincide. In that special case, the general hypothesis collapses to the ALM hypothesis. But in general, the routineness of a task is only one signal of a task’s routinisable. This is what recent literature exploring the new limits on automation have found (see, for instance, Brynjolfsson and Mitchell 2017; Brynjolfsson, Mitchell, and Rock 2018). In turn, and most importantly, the set of tasks that were routinisable in the past is likely to be far smaller than the set of tasks that are routinisable in the future.

4 Conclusion

The traditional task-based literature that explores the effect of technological change on the labour market, based on the ALM hypothesis, supports an optimistic view about the threat of automation. This optimism relies on the claim that there exists a large set of types of tasks that cannot be automated and, in turn, that those “tasks that cannot be substituted by automation are generally complemented by it” (Autor 2015). But if, as argued in this paper, many ‘non-routine’ tasks can also be automated, then the set of types of tasks that offer a ‘refuge’ for displaced labour will be far smaller than this task-based literature has traditionally assumed. Intuitively this suggests that the ALM hypothesis, and the firm boundary it has imposed between what systems and machines can and cannot do, may have created a misleading sense of optimism about the prospects for labour. This is an important argument for that requires further theoretical and empirical research. A first step towards this is an alternative hypothesis about the capabilities of new technologies – that is what I have set out in this paper.

5 Bibliography

- Acemoglu, Daron.** 2002. “Technical Change, Inequality, and the Labor Market”. *Journal of Economic Literature* 40 (1): 7-72.
- Acemoglu, Daron and David Autor.** 2011. “Skills, Tasks and Technologies: Implications for Employment and Earnings”. In: *Handbook of Labour Economics, Volume 4B* eds. David Card and Orley Ashenfelter, 1043-1171. Amsterdam: North-Holland.
- Acemoglu, Daron and Pascual Restrepo.** 2017. “The Race Between Machine and Man: Implications of Technology for Growth, Factor Shares, and Employment.” NBER Working Paper No. 22252.
- Acemoglu, Daron and Pascual Restrepo.** 2018. “Artificial Intelligence, Automation, and Work.” NBER Working Paper No. 24196.
- Autor, David.** 2013. “The “Task Approach” to Labor Markets: An Overview”, *Journal for Labour Market Research* 46 (3) 185 - 199.
- Autor, David.** 2014. “Polanyi’s Paradox and the Shape of Employment Growth”. NBER Working Paper No. 20485.
- Autor, David.** 2015. “Why Are There Still So Many Jobs? The History and Future of Workplace Automation”. *Journal of Economic Perspectives* 29 (3) 3-30.
- Autor, David and David Dorn.** 2013. “The Growth of Low-Skill Service Jobs and the Polarisation of the US Labor Market”. *The American Economic Review* 103 (5) 1553-1597.
- Autor, David, Frank Levy and Richard Murnane.** 2003. “The Skill Content of Recent Technological Change: An Empirical Exploration”. *The Quarterly Journal of Economics* 118 (4) 1279-1333.
- Autor, David and Anna Salomons.** 2017a. “Does Productivity Growth Threaten Employment? “Robocalypse Now?””. Presentation, European Central Bank Annual Conference, Sintra, Portugal, 27 June 2017.
- Autor, David and Anna Salomons.** 2017b. “Does Productivity Growth Threaten Employment? “Robocalypse Now?””. Paper for European Central Bank Forum on Central Banking.
- Brynjolfsson, Erik and Tom Mitchell** 2017. “What can machine learning do? Workforce implications”. *Science* 358 (6730) 1530-1534.
- Brynjolfsson, Erik, Tom Mitchell and Daniel Rock** 2018. “What Can Machines Learn, and What Does it Mean for Occupations and the Economy?”. *American Economic Review Papers and Proceedings* 108 43-47.
- Caselli, Francesco and Alan Manning.** 2018 (forthcoming). “Robot arithmetic: new technology and wages”. *The American Economic Review: Insights*.
- Dreyfus, Hubert.** 1976. *What Computers Can’t Do: The Limits of Artificial Intelligence* New York: Harper & Row.
- Dreyfus, Hubert.** 1992. *What Computers Still Can’t Do: A Critique of Artificial Reason* MA: MIT Press.
- Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan Swetter, Helen Blau and Sebastian Thrun.** 2017. “Dermatologist-level classification of skin cancer with deep neural networks.” *Nature* 542 155-188.
- Gibbs, Samuel.** 2016. “Self-driving cars: who’s building them and how do they work?” *The Guardian*, May 26.
- Goos, Maarten, Alan Manning and Anna Salomons.** 2014. “Explaining Job Polarization: Routine-Biased Technological Change and Offshoring”. *The American Economic Review* 104 (8) 2509-2526.
- Hayes-Roth, Frederick, Donald Waterman and Douglas Lenat.** 1983. *Building Expert Systems*. MA: Addison-Wesley.

- ILO.** 2016. ‘New technologies: A jobless future or golden age of job creation?’, Working Paper No. 13.
- IMF.** 2017. *World Economic Outlook*.
- Kasparov, Garry.** 2010. “The Chess Master and the Computer” The New York Review of Books, Feb 11.
- Levy, Frank and Richard Murnane.** 2004. *The New Division of Labor*. India: New Age International Ltd.
- Neumann, Dana.** 2016. “Human Assistant Robotics in Japan’ EU Japan Centre for Industrial Cooperation, March.
- Ng, Andrew.** 2016. “What AI Can and Can’t Do Right Now’ The Harvard Business Review, November 9.
- Polanyi, Michael** 1966. *The Tacit Dimension*. London: University of Chicago Press.
- Puzder, Andy.** 2016. “Why Restaurant Automation Is on the Menu” The Wall Street Journal, March 24.
- Remus, Dana and Frank Levy.** 2016. “Can Robots be Lawyers? Computers, Lawyers, and the Practice of Law.” Available at SSRN: <http://ssrn.com/abstract=2701092>.
- Russell, Stuart and Peter Norvig** 1995. *Artificial Intelligence: A Modern Approach*. New Jersey: Pearson Education.
- Selbst, Andrew and Julia Powles** 2017. “Meaningful information and the right to explanation”. *International Data Privacy Law* 7 (4) 233-242.
- Silver, David, Aja Huang, Chris Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel and Demis Hassabis.** 2016. “Mastering the game of Go with deep neural networks and tree search.” *Nature* 529 484-489.
- Weizenbaum, Joseph.** 1976. *Computer Power and Human Reason*. US: W.H. Freeman and Company.
- The White House.** 2016. ‘Artificial Intelligence, Automation, and the Economy’.
- Winston, Patrick.** 1977. *Artificial Intelligence*. Reading: Addison Wesley.
- The World Bank.** 2016. *World Development Report 2016*.