Firms as problem solvers: economics meets computer science

Kim, Minseong

University of Illinois at Urbana-Champaign

1 December 2019

Firms as problem solvers: economics meets computer science

By Minseong Kim*

A theory of the firm based on the idea that firms are problem solvers is developed. A network of firms and hierarchical structure of an individual firm are analyzed in terms of distributed and parallel computing. This framework, based on notions of computer science, allows us a simple answer to why a network of firms exists instead of a network of individuals.

JEL: D20, D40, L20

Keywords: theory of the firm, vertical integration, parallel computing, distributed computing, synchronization

I. Introduction

As noted in Coase (1937) and Simon (1991), a modern economy is much closer to networks of firms rather than to networks of individuals. Each firm is marked by central planning rather than a decentralized market. This is a strange state of affair in a neoclassical theory: how is it that decentralization works well between firms and yet somehow advantages of decentralization are not fully exploited inside each firm? What makes a firm particular different from a simple network of individuals? That is, why does a firm governed with some degrees of centralization even exist?

The field of 'the theory of the firm' tries to answer this question. Proposed

* Minseong Kim: University of Illinois Urbana-Champaign, mkim146@illinois.edu

theories are numerous: transaction cost theory (Coase, 1937; Williamson, 1971; Klein, Crawford and Alchian, 1978), residual control rights theory (Grossman and Hart, 1986), knowledge-based view (Kogut and Zander, 1992), agency theory (Holmstrom and Milgrom, 1991) and so on. These views tend to rely on incomplete contracts, market incompleteness or immobility of some factors. Maskin and Tirole (1999) offers a criticism on these views - if contract incompleteness is due to transaction costs of describing or forecasting future contingencies, then the issues can be made irrelevant. Responses to the Maskin-Tirole critique have since appeared (Hart and Moore, 2008; Aghion et al., 2012).

Where we depart from usual conventions is this: an economic process is constrained by physics and computer science constraints. The latter comes to primary importance, when we consider firms as trying to solve some problems. Given the Church-Turing thesis (Rosser, 1939), solving a problem amounts to computation that results in a solution. Thus computer-scientific notions come to matter.

Whether a firm is simply executing a well-crafted production plan, researching for a new product, or responding to some consumer demands, it is computing something so that one gets a solution, which becomes a product, service or something else.

Given the above, we may attempt to explain what a firm is and why firms form by references to notions of computer science. In particular, notions used in parallel and distributed computing will be particularly useful. We will not rely on contract incompleteness or the idea that some factors, such as know-hows, are hard to replicate. We will derive a theory of the firm purely based on considerations of what agents face when they try to solve

a problem.

The framework we develop in this paper is appealingly simple relative to other common frameworks and complementary to conventional approaches. We will see that the computer-scientific framework allows us to address the criticism commonly made against a transaction costs theory: that the line between bureaucratic costs and transaction costs is blurry as to provide an account of why a firm exists. The key notions of computer science used in this paper are those of parallel computing (Padua, 2011) and distributed computing (van Steen and Tanenbaum, 2016).

We use 'distributed system' and 'distributed network' interchangably in this paper.

II. The theory of the firm: parallel versus sequential executions

A. Notions of the theory of the firm in distributed and parallel computing terminology

- We associate a firm with a central processing unit (CPU).
- A CPU can be multi-core: it consists of multiple processing units - cores. Each division or department of a firm is associated with each core.
- In this paper, a computer will be considered indistinguishable from a CPU, as distinguishing them is irrelevant for purposes of this paper.
- Markets of firms and consumers are considered to be distributed networks of computers.

We now can note the following: since markets are considered to be distributed networks, results in distributed computing are relevant. For a firm,

results in parallel computing are relevant.

B. *Parallel computing*

A problem takes time to solve. Thus, one of central questions in computer science is about time complexity - how can we reduce time in solving a problem? If we can parallelize solving a problem - that is, one core solves a sub-problem of the problem while another core solves another sub-problem that does not rely on a solution of the first sub-problem and then later combine these results, then we may get speed-ups.

In terms of a firm, this involves somewhat autonomous departments or divisions. Each department works to solve its own problem with some independence from other departments and upper hierarchical commands, later to combine outputs with other departments.

However, one cannot forever divide a problem into sub-problems - there is limitation to how far parallelization can go. Inherently sequential problems that cannot be parallelized exist. Solving a problem then requires sequential execution - things have to be executed step-by-step sequentially.

C. *Department as an irreducible sequential unit: why central planning is at least minimally required*

By the above, we can define a department or division of a firm as an irreducible sequential unit that cannot be parallelized.

While each worker of a department has their own individual task, workers must follow the shared command so that a sequential step can properly executed.

In synchronous CPU terminology, at each clock cycle, each component of core must execute each step as given by command signals. Asynchronous

CPU designs make this story somewhat complicated, but the fact that at least some minimal levels of commands and central planning are required in order for computation to work remains.

Thus by recourse to computer-scientific constraints, we can explain why central planning is somewhat necessary. But central planning is bad when there can be further parallelization - it is better to split bureaucratic structures so that a problem can be solved more efficiently.

D. When does a unit becomes a department, and when does a unit becomes a separate firm?

A firm is a distributed network of departments. Similarly, a market is a distributed network of firms. While we answered why a separate department consisting of individuals exists, we have not answered why a firm consists of multiple departments, instead of an individual department being a firm itself.

This is in essence the common criticism against a transaction costs theory translated: there are costs to communications both between departments and between firms. So why does not each department become a firm on its own?

There are two parts in answering the question.

- Parallelizing an algorithm for solving a problem does not completely eliminate the question of command. At some points, inputs to and outputs of each core have to be processed and coordinated by some central commands.

- A sub-problem that each core solves may only have relevance for a particular problem. That is, a sub-problem is not marketable. This is

identical to the circumstance of labor monopsony: a particular firm is the dominant buyer of labor power of workers in a division.

In reality, workers can often re-train themselves to do other works, thus scale of labor monopsony is constrained. Note that there are some evidences for significant labor monopsony in labor markets. (Dube et al., n.d.)

Existence of a firm, however, does not depend on existence of significant labor monopsony: the point here is that if a firm solves a particular problem, then parallel sub-problems that each department solves often only have meaning for that particular problem. In such a case, each department is almost always subject to central commands of a firm: it is dictated by upper hierarchies to solve a particular sub-problem with a particular input or a goal. We need both above points to get the proper theory of the firm.

III. The theory of the firm: synchronous versus asynchronous networks

A. What about vertical integration?

Different parts of a supply chain can be outsourced to other firms. In our framework, this would be about different firms solving separate sub-problems that form a bigger problem of supplying a product or a service. Some of these sub-problems often are 'marketable' in sense that they are invoked by other big problems as well.

So in such a case, why does a firm sometimes choose to integrate different parts of a supply chain? For this, FLP theorem (Fischer, Lynch and Paterson, 1985) in distributed computing becomes important.

B. *Asynchronous consensus problem*

Economic coordination is a specialized form of consensus - agents agree to do something. We now initially think of a network of firms as an asynchronous distributed system. The main feature of an asynchronous distributed system is that there is no upper bound on single message delivery delay duration. Suppose realistically that at least one processing unit (firm) may fail. Note that the word is 'may fail' not 'fail'. Thus there may actually be no processing unit that failed.

Then FLP theorem states that there exists no deterministic consensus protocol that ensures agreement, termination and validity. Roughly translated, agreement means that every consensus-arriving unit agrees on same consensus, termination means that all units that do not fail eventually reach consensus, validity is that consensus is what would have been intended - in other words, units cannot simply declare do-nothing as consensus.

Loopholes exist, however with some limitations. (Aspnes, 2003)

- Randomization. we can adopt a randomized consensus protocol instead of a deterministic consensus protocol. However, a randomized consensus protocol still allows for non-consensus results - it is just that they collectively have probability of 0. While given infinite time, it is almost as if they rarely occur, they still occur.
- Make the network at least partially synchronous.
- Failure detectors. This is also about how an asynchronous system approximation to real networks must be 'strengthened' as to allow consensus.
- 'Shared memory' strengthening of asynchronous systems.

All of these, except for randomization, effectively are about obtaining a partially synchronous system in more general sense, and having synchrony means that there is upper bound to message delay duration.

In case of firms, if firm A does not get response from firm B in some pre-set durations, then A assumes that B failed, instead of having delays in delivering its message which might be the case. A buyer firm of some product or service can fire a contractor if it does not respond in some given time. Roughly, this works as a very crude failure detector that has some degree of inaccuracy in detecting actual failures.

Given these necessities of reporting for coordination or common infrastructure ('shared memory'), a firm may decide that it is better to eliminate double bureaucratic inefficiencies and simply have one central command unit. That is, because there are costs to synchronization when problem solving is parallelized, for some problems more sequential algorithms are better in efficiency.

C. Bureaucratic costs and transaction costs: sequentialization versus parallelization

We thus evaded the concern typical in the transaction costs theory: that bureaucratic costs and transaction costs are largely similar.

Bureaucratic costs are about costs of sequentialization. Transaction costs are about costs of parallelization. The size of a firm is determined by optimizing against these costs.

While in a traditional economics understanding, these may be just word plays, in a computer-scientific understanding, sequentialization and parallelization have precise meanings.

IV. Conclusions

Let us summarize. The theory of the firm is about explaining why a firm, governed with central planning, exists instead of an efficient market network of individuals.

- A firm is understood as a solver of a particular problem.
- Bureaucratic costs are costs of sequential approaches to solving a problem. Transaction costs are costs of parallel approaches to solving a problem, given necessities of some degree of synchronization. Unlike in a traditional economics understanding, they have precise meanings and distinctions in a computer science understanding.
- There is no need for contract incompleteness to justify existence of a firm.
- A department or a division of a firm exists because a problem that a firm tries to solve may be partially broken down into sub-problems that can solved in a parallel way: other departments solve other sub-problems, while 'my' department solves a particular sub-problem simultaneously.
- A department does not become a firm itself because despite parallelism, almost no problem can be made perfectly parallelized and thus central commands are necessary. If a sub-problem that each department solves is marketable in sense that other problems require solving the sub-problem, then a department can break out to become a new firm.
- A department is a minimal sequential (processing) unit, while a firm

is a collection of sequential units heavily geared toward solving a particular problem.

- The key to the theory of the firm in this paper thus is sequentialism versus parallelism. Under this background, the question of asynchronous versus synchronous networks is important in explaining why different firms integrate into one.

REFERENCES

Aghion, Philippe, Drew Fudenberg, Richard Holden, Takashi Kunimoto, and Olivier Tercieux. 2012. "Subgame-Perfect Implementation Under Information Perturbations*." *The Quarterly Journal of Economics*, 127(4): 1843–1881.

Aspnes, James. 2003. "Randomized Protocols for Asynchronous Consensus." *Distributed Computing*, 16(2-3): 165–175.

Coase, R. H. 1937. "The Nature of the Firm." *Economica*, 4(16): 386–405.

Dube, Arindrajit, Jeff Jacobs, Suresh Naidu, and Siddharth Suri. n.d.. "Monopsony in Online Labor Markets." *American Economic Review: Insights*.

Fischer, Michael J., Nancy A. Lynch, and Michael S. Paterson. 1985. "Impossibility of Distributed Consensus with One Faulty Process." *J. ACM*, 32(2): 374–382.

Grossman, Sanford J., and Oliver D. Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy*, 94(4): 691–719.

Hart, Oliver, and John Moore. 2008. "Contracts as Reference Points*." *The Quarterly Journal of Economics*, 123(1): 1–48.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization*, 7: 24–52.

Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian. 1978. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." *The Journal of Law and Economics*, 21(2): 297–326.

Kogut, Bruce, and Udo Zander. 1992. "Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology." *Organization Science*, 3(3): 383–397.

Maskin, Eric, and Jean Tirole. 1999. "Unforeseen Contingencies and Incomplete Contracts." *The Review of Economic Studies*, 66(1): 83–114.

Padua, David. 2011. *Encyclopedia of Parallel Computing.* Springer Publishing Company, Incorporated.

Rosser, Barkley. 1939. "An informal exposition of proofs of Gödel's theorems and Church's theorem." *Journal of Symbolic Logic*, 4(2): 53–60.

Simon, Herbert A. 1991. "Organizations and Markets." *Journal of Economic Perspectives*, 5(2): 25–44.

van Steen, Maarten, and Andrew S. Tanenbaum. 2016. "A brief introduction to distributed systems." *Computing*, 98(10): 967–1009.

Williamson, Oliver E. 1971. "The Vertical Integration of Production: Market Failure Considerations." *The American Economic Review*, 61(2): 112–123.