Spring 2020

Internal Badger Workshop

Session 8: Economies of Badger

Moderator: Marek Rosa

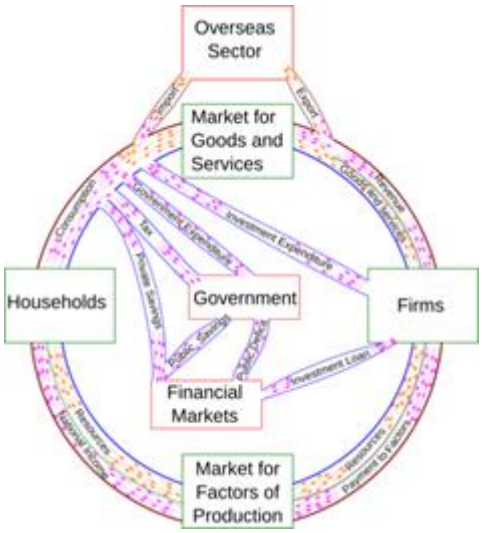
Pre-discussion Comments & Resources

Standard definitions

- An economy is an area of the production, distribution and trade, as well as consumption of goods and services by different agents.
- Economics is the social science that studies the production, distribution, and consumption of goods and services.

Badger specific

- Economics in Badger is a study of information exchange in the network of experts where the information can be interpreted as a flow of credit, from the environment to the computational resources.



Why?

- Network of experts face the "credit assignment problem" in situations in which only incomplete performance evaluations are available. The credit assignment problem is that a network should assign credit or blame for its behaviours according to the contribution to the network performance.
- Economy helps solve this problem in a distributed and localized manner
- Joseph Davidson:
    - ...in cooperative tasks, optimizing for your individual reward is better than optimizing for the global one. The reason being that individual rewards provide feedback into specifically how you are doing, which provides the most information for improvement. If there is a global reward that everyone contributes to, your performance isn't as easy to determine as it is now the average of everyone, including those that do great and those that do awfully.
    - Individual reward being more informative than the global reward.
    - Conclusion: It seems that it explains why the capitalism (market economy) was more efficient than communist' centralized planning.

Ideas

- Selfish experts, each one maximizing its own profit (the difference between credit and cost)
- Experts dissipate unless they "pay" credit to the computational resources (cost)
- Expert compete for credit
- Experts can outsource some tasks to other experts, or hierarchies of experts. Experts pay for this service (credit distribution)
- Experts learn how to evaluate the contribution of individual experts towards a higher level task
- Localized and distributed credit assignment is continuously being learnt
- Credit assignment is adaptable and not hardcoded
    - Example: MetaGenRL - learned loss function, learned credit assignment. What if we make it more modular and open ended?
- Credit signals allow the experts to coordinate in real time, and therefore to adapt to change in a way that a fixed credit assignment could never allow.
- Credit comes from outside of the agent, it's an unlimited quantity
- Credit dissipates from agent, consumable limited quantities
- Economy solves the competition for limited resources, is adaptable to novel problems

Opinion

- On an individual level, experts are competing between each other for limited computational resources.
- On a collective level, the society of experts (the agent) is competing with the external environment for collective survival.
- If experts share limited resources inefficiently, they decrease the survivability of the entire collective.

Alternative view

- Credit is just an online feedback from the environment, for real-time adaptation, allocating limited resources where they are needed most
- Payments to computational resources are just a form of prioritization of the limited computational resources usage.
- There is no credit income
- There is no credit dissipation
- Question: How is this flow different from our universe?

Alternative view 2

- Energy flows through an ecosystem in only one direction. Energy is passed from organisms at one trophic level or energy level to organisms in the next trophic level.
- Contructal law says that flow systems will shape themselves to flow more easily over time, eventually making a tree shape - Adrian Bejan=

Open questions

1. Economy as a dissipative system? [Shiozawa 1996]
2. How complex economy is too much? Brain neurons are not very complex but perhaps neural columns could enable more complex economic institutions and credit interchange?
3. Which of these economic phenomena can be expected in future Badger prototypes?
    a. Different currencies (different types of credit)
    b. Specialization
    c. Inflation
    d. Monopoly on intellectual property, versus, copy left intellectual property. What would be the impact on agent's growth?
    e. Auctions
    f. Stock markets, money lending, startups, VC funds, Private Equity
4. How to test for them?
5. Which phenomena are certainly going to be present (e.g. different currencies) and which are unlikely (e.g. stock market, money lending, money laundering)
6. Is it even useful for our research to try to seek economical analogies in Badger?
7. Maybe the external credit and payment to computational resource is not necessary. What if the credit should be an internal property, an estimation of how different experts contribute to agent's survival and the payment to the computational resources is just a necessity?
8. Is Badger an open or closed thermodynamic system?

Next steps

1. Agent-based model prototype visualizing the flow of credit from external world and dissipated in the computational resources

Discussion Notes

- Economy shows how multi-agentness can be useful - instead of one single loss, you can have multiple independent (incompatible, not shared) losses
- Population based optimization (?)
- Experts should not be printing their own credit as it would not bring any benefit. The question is that if we train the expert policy via the outer loop, how can we prevent discovery of a policy that will heavily rely on printing money? In other words, we should have a strong structure that will prevent printing money, otherwise it's hard to get rid of this kind of cheating.
- Selfish experts each one maximizing its own profit (credit - cost) is a nice solution to expert diversification (heterogeneous behaviour, breaking the symmetries).
- There seems to be a link between the concept of an economy and the meta-reasoning stage in badger, especially when the concept of currency is used for "paying" for compute
- The concept of an economy is beneficial for multi-agentness
    - One reason is the ability to naturally have multiple loss functions
    - This means there is no fixed point
    - In a way the system then acts like a game, e.g. ala GAN, where there is a continual interplay between the losses
    - This can have many benefits, but also makes things harder to train
    - It is not straightforward to have a single loss in a way that would match behaviour of multiple losses
    - Dynamics of the system can be richer with multiple losses than with a single loss
- Game Theory vs Economy
    - Viewpoint of Ecology might also be beneficial, despite both sharing a lot
- Question: What's the primary reason for having an economy?
    - Dealing with limited computational Resources / assignment of them
    - A link between the local and the global behaviour
        - If we assume that most interactions are local, economy and the exchange of credit will all the feedback (in both directions) between the micro and the macro level and hence influence between the two
        - A substrate for global influence
- Economy shut's down same experts
    - We might need a Lagrangian (Soft) death to make things learnable using gradient descent
- Maximising your own reward selfishly (as in the IC3Net paper) may or may not have some asterisks next to it, as the idea for selfish rewards maximised trajectories which would be beneficial to the global reward if all agents were to adopt that strategy - but we would have to be careful about how orthogonal the global rewards and individual rewards are.
    - In their traffic junction example: A car driving well helps everyone, the agent controlling the car gets a good individual reward for that. Conversely a bad car (can) hurt everyone, so the trajectory for that is individually punished. This is better than averaging the good and bad trajectories for a global reward.
    - In section 4.3 of the IC3Net paper, they state that the IC3Net credit assignment with individual rewards shows better convergence than vanilla CommNet with global rewards

    - A selfish reward scheme can be used to bootstrap an expert population, then shifting over to a more global goal.
        - Broadly mirroring human development from caring about only yourself -> caring about family -> caring about neighbours -> community -> country -> globe
    - If the individual reward is too orthogonal, then cooperation will be seen as strictly bad relative to the objective measure and cooperation will never emerge.

References/Researchers mentioned during discussion

- Eric B. Baum, Toward a Model of Intelligence as an Economy of Agents, 1999
- Letcher et al, Stable Opponent Shaping in Differentiable Games, ICLR 2019
- Wang et al., Shapley Q-value: A Local Reward Approach to Solve Global Reward Games, AAAI 2020
- Sing, Jain and Sukhbaatar, Learning when to Communicate at Scale in Multiagent Cooperative and Competitive Tasks, ICLR 2019
- Berseth et al., SMiRL: Surprise Minimizing RL in Dynamic Environments, arXiv 2019
    - just preservation is enough to learn interesting things
- Leveraging Environmental Correlations: The Thermodynamics of Requisite Variety
    - Other works of Jim Crutchfield, Cosma Shalizi and Adrian Bejan
- Literature on *n*-auction design, incentive design, market design. Quadratic funding is interesting for example.