Decentralized Swaps Problems in Blockchains

Marianna Belotti[1] , Maria Potop-Butucaru [2] , Stefano Moretti[3] and Stefano Secci[4]

LINCS Paris, France

12/06/2019

---

[1]Groupe Caisse des Dépôts - Cnam

[2]Sorbonne Université

[3]Université Paris Dauphine

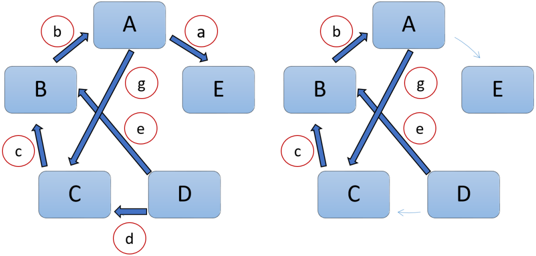
[4]Cnam

What is a swap ?

A *swap situation* is a tuple $\langle \mathcal{A}, \mathcal{O}, b_0, b_*, (u_i)_{i \in \mathcal{O}} \rangle$ where:

- $\mathcal{A} = \{1, \dots, m\}$ is the set of *assets*;
- $\mathcal{O} = \{1, \dots, n\}$ is the set of *owners* or *agents*, with $m \geq n$;
- $b_0, b_* : \mathcal{A} \to \mathcal{O}$ (both **surjective**) the *original* and the *desired* ownership map, respectively;

What is a swap ?



- u_i is the payoff function for owner $i \in \mathcal{O}$ over bundles of assets in $2^{\mathcal{A}}$ such that $u_i(b_0^{-1}(i)) < u_i(b_*^{-1}(i))$ and for any $S, T \in 2^{\mathcal{A}}$ with $S \subseteq T$ we have $u_i(T) \geq u_i(S)$, for each $i \in \mathcal{O}$.

Multi-Swap Protocols

1. A *multi-swap protocol* is a sequence of pairs (transfers) $\sigma = \{(a_i^k, o_j^k)\}_{k\in\{1,2,\ldots,t\}}$ with $i \geq j,\ a_i \in P(\mathcal{A})$ and $o_j \in P(\mathcal{O})$, where $P(\mathcal{A})$ and $P(\mathcal{O})$. A pair (a_i^k, o_j^k) has the meaning that assets in a_i are transferred at step k to owners o_j.
2. A *single-swap protocol* consists in a sequence $\sigma = (a^1, o^1), (a^2, o^2), \ldots, (a^t, o^t)$ where $|o^k| = |a^k| = 1$, for all $k \in \{1, 2, \ldots, t\}$. A single-swap protocol engenders a sequence of maps $b_1^\sigma, b_2^\sigma, \ldots, b_t^\sigma : \mathcal{A} \to \mathcal{O}$ such that for all $k = 1, \ldots, t$:
   - $b_k^\sigma(z) = b_{k-1}^\sigma(z)$ for all $z \in \mathcal{A} \setminus \{a^k\}$;
   - $b_k^\sigma(a^k) = o^k$.

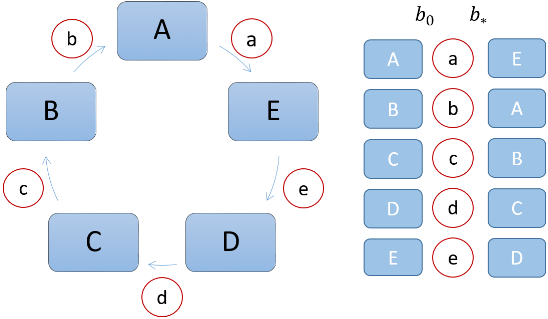
Multi-Swap Protocols

1. A *multi-swap protocol* is a sequence of pairs (transfers) $\sigma = \{(a_i^k, o_j^k)\}_{k \in \{1,2,\dots,t\}}$ with $i \geq j,\ a_i \in P(\mathcal{A})$ and $o_j \in P(\mathcal{O})$, where $P(\mathcal{A})$ and $P(\mathcal{O})$. A pair (a_i^k, o_j^k) has the meaning that assets in a_i are transferred at step k to owners o_j.

2. A *single-swap protocol* consists in a sequence
$$\boxed{\sigma = (a^1, o^1), (a^2, o^2), \dots, (a^t, o^t)}$$
where $|o^k| = |a^k| = 1$, for all $k \in \{1, 2, \dots, t\}$. A single-swap protocol engenders a sequence of maps $b_1^\sigma, b_2^\sigma, \dots, b_t^\sigma : \mathcal{A} \to \mathcal{O}$ such that for all $k = 1, \dots, t$:
   - $b_k^\sigma(z) = b_{k-1}^\sigma(z)$ for all $z \in \mathcal{A} \setminus \{a^k\}$;
   - $b_k^\sigma(a^k) = o^k$.

Correspondence with graphical representation

From the information contained in the tuple $\langle \mathcal{A}, \mathcal{O}, b_0, b_*, (u_i)_{i \in \mathcal{O}} \rangle$ we can construct a digraph $D = (E, V)$.



Proposition

Consider a protocol $\sigma = (a^1, o^1), (a^2, o^2), \ldots, (a^t, o^t)$. Then replacing o^k by $b^{\sigma}_{k-1}(a^k)$ in σ, with $k \in \{1, \ldots, t\}$ (i.e., we consider a new sequence $\sigma^k = (a^1, o^1), (a^2, o^2), \ldots, (a^{k-1}, o^{k-1}), (a^k, b^{\sigma}_{k-1}(a^k)), (a^{k+1}, o^{k+1}), \ldots, (a^t, o^t)$, for some $k \in \{1, \ldots, t\}$) implies that:

(i) $(b^{\sigma^k}_t)^{-1}(o^k) \subseteq (b^{\sigma}_t)^{-1}(o^k)$ and,

(ii) $(b^{\sigma^k}_t)^{-1}(b^{\sigma}_{k-1}(a^k)) \supseteq (b^{\sigma}_t)^{-1}(b^{\sigma}_{k-1}(a^k))$.

Decision Function

Definition

A *decision function* as a map $F : \{1, \ldots, t\} \to \mathcal{O}$ that specifies which owner $F(k)$ has the power to decide at step k whether to transfer a^k to o^k (i.e., follow the protocol) or to leave it to the owner of a^k at step $k - 1$.

Definition

A decision function F is ***effective*** on σ if $F(k) = o^k$ for any $k \in \{1, \ldots, t\}$ (so, agent o_k has the power to accept or not asset a^k).

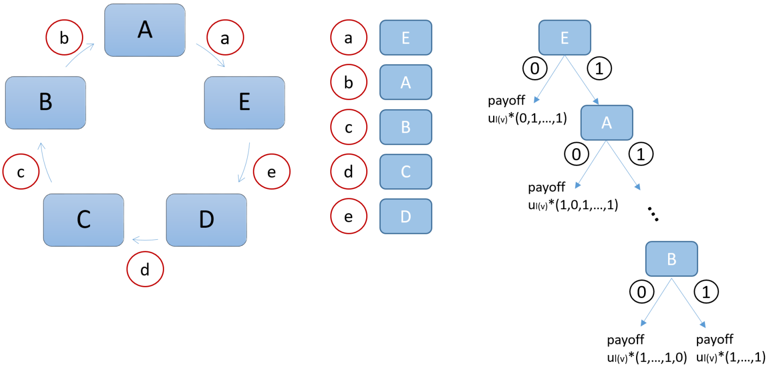
Associated Extensive form Game

Definition

Let $\langle \mathcal{A}, \mathcal{O}, b_0, b_*, (u_i)_{i\in\mathcal{O}} \rangle$ be a swap, $\sigma = (a^1, o^1), (a^2, o^2), \ldots, (a^t, o^t)$ be a single-swap protocol and F be a decision function. We define the extensive game form $\Gamma^\sigma = \langle \mathcal{O}, T, P, (A_h)_{h\in V}, (u_i)_{i\in N} \rangle$ such that:

- T is a (binary) directed tree such that each directed path from the root v_0 to an end node $v \in Z$;
- $P(v) = F(l(v))$, for each $v \in V \setminus Z$ is the activator at step $l(v)$ in the protocol σ, where $l(v)$ is the number of arcs between v_0 and v;
- A_h for all $h \in V$ is formed by two outgoing arcs in h; one arc in A_h is labeled with **action 1** (i.e., follow the protocol σ) and the other one with label **0** (i.e., not follow the protocol σ) for any $h \in V \setminus Z$.
- Any end node $z \in Z$ is associated to a unique outcome corresponding to $b_t^{\sigma^K}$ where $K \subseteq \{1, \ldots, t\}$ is such that $p_k^z = 0\ \forall k \in K$, and $p_k^z = 1$ for any $\{1, \ldots, t\} \setminus K$. So, for any $i \in \mathcal{O}$, the outcome $b_t^{\sigma^K}$ is evaluated by i with the payoff function $u_i((b_t^{\sigma^K})^{-1}(i))$.

Associated Extensive form Game



Subgame perfect equilibrium

Proposition

Let $\Gamma^\sigma = \langle N, T, P, (A_h)_{h \in V}, (u_i)_{i \in N} \rangle$ be the extensive form game associated to a swap situation $\langle \mathcal{A}, \mathcal{O}, b_0, b_*, (\succeq_{u_i})_{i \in \mathcal{O}} \rangle$, a single -swap protocol $\sigma = (a^1, o^1), (a^2, o^2), \ldots, (a^t, o^t)$ and let $F : \{1, \ldots, t\} \to \mathcal{O}$ be a decision function. If F is **effective** on σ, then the strategy profile $(\hat{s}_1, \ldots, \hat{s}_n)$ that specifies action 1 at any node is the unique subgame perfect equilibrium (in dominant strategies).

For each node v, by the fact that F is effective, we have that $P(v) = F(l(v)) = o^{l(v)}$. So, at each decision node $v \in V \setminus Z$, if player $P(v)$ specifies action 0 at node v, then by the first claim of Proposition 1, player $P(v)$ ends up with a set of assets that is contained in the one that player $P(v)$ would obtain if she/he specifies action 1 at node v.

Subgame perfect equilibrium

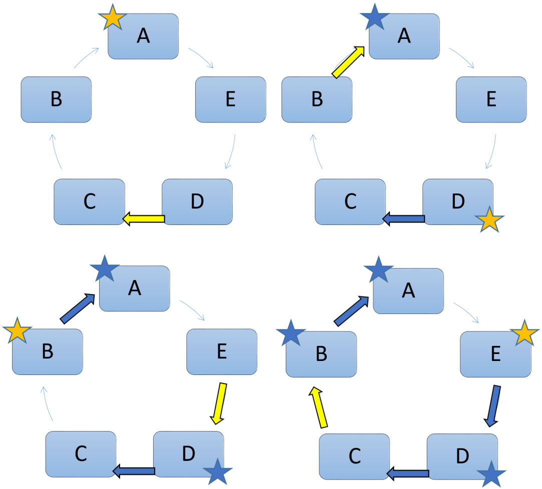
Proposition

The decision function $F : \{1, \dots, t\} \rightarrow \mathcal{O}$ is effective on single-swap protocol $\sigma = (a^1, o^1), (a^2, o^2), \dots, (a^t, o^t)$, if and only if the strategy profile $(\hat{s}_1, \dots, \hat{s}_n)$ that specifies action 1 at any node is the unique subgame perfect equilibrium (in dominant strategies).

The "*if*" has to proved since the "*only if*" follows from Proposition 2. **By contradiction**, if F is not effective we have the following cases: (i) $P(v) = F(l(v)) = b^{\sigma}_{l(v)-1}(a^{l(v)})$, the original owner of the asset decide whether or not to follow the protocol σ or, (ii) $P(v) = F(l(v)) = o^j \neq b^{\sigma}_{l(v)-1}(a^{l(v)})$ such that $j \neq l(v)$, the activator is any player in the game but the original asset owner and the asset receiver.

Case (i) derives from Proposition 1.

Case (ii) more complex.



Solution: Penalty mechanism

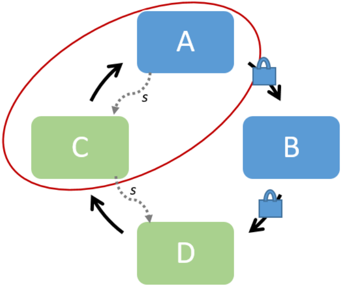
Proposition

If the decision function $F : \{1, \dots, t\} \to \mathcal{O}$ is not effective on single-swap protocol $\sigma = (a^1, o^1), (a^2, o^2), \dots, (a^t, o^t)$, then the strategy profile $(\hat{s}_1, \dots, \hat{s}_n)$ that specifies action 1 at any node is the unique subgame perfect equilibrium (in dominant strategies) only with a penalty function $p : A_h \to \mathbb{R}_+$ constructed as following:

- $p(0; b^{\sigma}_{l(v)-1}(a^{l(v)})) = u_{o^{l(v)}}\left((b^{\sigma}_t)^{-1}(o^{l(v)})\right) - u_{o^{l(v)}}\left((b^{\sigma^j}_t)^{-1}(o^{l(v)})\right)$ and,
- $p(0; j) = \epsilon \in \mathbb{R}_+$ for $j \neq b^{\sigma}_{l(v)-1}(a^{l(v)})$.

Cooperative Approach for HTL protocol

The aim of using cooperative game theory is to analyze the coalition formation process in the case where agents aim at **taking advantage** of the *atomic cross-chain swap protocol* ending up with a possible gain. Therefore, we concentrate on a **optimistic** vision of the game comparing the best outcomes coalitions can reach.



Strategies

Strategies:

- **Follow**: each player follow the HTL protocol in every step from the contract publication on the blockchain with the hashlock to the secret revelation to the correct player.
- **Deviate**: the player decide to behave *irrationally* or *maliciously* and decide not to publish or not to trigger the blockchain transaction.

Outcomes

The possible outcomes for player $o^k \in \mathcal{O}$:

1. **Free Ride (FR)**: the player acquires assets without paying; $o^k \in \mathcal{O}$ ends up with a free ride whenever $b_0^{-1}(o^k) \subsetneq (b_t^\sigma)^{-1}(o^k)$.
2. **Discount (DS)**: the player acquires assets while paying less than expected; $o^k \in \mathcal{O}$ has a discount whenever $b_*^{-1}(o^k) \subsetneq (b_t^\sigma)^{-1}(o^k)$.
3. **No Deal (ND)**: the player's asset(s) does(do) not change hands; $o^k \in \mathcal{O}$ ends up with a no deal whenever $b_0^{-1}(o^k) = (b_t^\sigma)^{-1}(o^k)$.
4. **Deal (D)**: the player swaps assets as expected. Player $o^k \in \mathcal{O}$ ends up with a deal whenever $b_*^{-1}(o^k) = (b_t^\sigma)^{-1}(o^k)$.
5. **Underwater (U)**: the player pays without acquiring all expected assets. Player $o^k \in \mathcal{O}$ goes underwater whenever $(b_t^\sigma)^{-1}(o^k) \subsetneq b_0^{-1}(o^k) \vee (b_t^\sigma)^{-1}(o^k) \subsetneq b_*^{-1}(o^k)$.

Hedonic Games

Hedonic games consider coalition formation in an environment where each players payoff is completely determined by the identity of other members of his coalition (hedonic setting).

Same preference profile on the outcomes for all the players:

- **D** $>$ **ND**: each player prefers an agreement because otherwise it would have her funds blocked for a time;
- **FR** $>$ **ND**: because it acquires additional assets.
- **DS** $>$ **D**: each player evidently prefers a discount to deal.
- **ND** $>$ **U**: each player does not like to loose money.

Individual Stability

Definition

A coalition partition P is *individually stable* if there do not exist $i \in N$ and a coalition $C_k \in P \cup \emptyset$ such that $C_k \cup \{i\} \succ_i C_P(i)$, and $C_k \cup \{i\} \succeq_j C_k$ for all $j \in C_k$.

Theorem

The partition consisting in the solo leader l and the followers all together is **individually stable**.