Linguo Parametrization

William George

These notes briefly describe the modifications to the general curated list model presented in https://github.com/kleros/research-docs/blob/master/klerosfees12fev2019.pdf that are required for Linguo parametrization. Note that the auction mechanism of Linguo pricing is such that different prices can be selected for tasks. Then the same contract, and hence the same parameters, such be acceptable for these different values. In some sense, for the purposes of this model, Linguo is an infinite collection of curated lists of translations, one for each price point.

Taking gas estimates from here: https://docs.google.com/spreadsheets/d/1m7Lv-WJiKwBGFcx0F--_wOwz6rWOZRubAOK8NCB18kQ/edit#gid=0 the gas required to challenge a submission is much larger than the gas required for any of the various other actions performed by the translator and challenger. (It is somewhat larger than the gas required by the requestor to create a task, however the requestor's incentives are less present in this model. We consider the game between translators/submitters and challengers, where attackers are translators that submit spam/false translations.) Hence, we model the challenger gas as a variable, but for simplicity we consider translator gas to be zero.

We take a slightly modified version of the notation of https://github.com/kleros/research-docs/blob/master/klerosfees12fev2019.pdf

- x price of task; note that this is the value of a successful attack
- e effort required to evaluate task - we model this as scaling with the price, so $e = e_0 x$ for $e_0 \in \mathbb{R}_{\geq 0}$
- p probability that a participant who puts in effort e correctly determines whether a translation is acceptable
- F arbitration fees
- g gas to challenge an item
- S_C challenger reward in our model this is of the form $nF + mx$, for $n, m \in \mathbb{R}_{\geq 0}$; correspondingly the translator deposit is $(n+1)F + mx$
- K number of active challengers
- y percentage of cases evaluated by each challenger
- u percentage of list consisting of challengeable submissions

Note that K, y, and u are arrived at in some equilibrium of the submitter/challenger game.

The attacker payoff for making a hostile submission is essentially unchanged from `https://github.com/kleros/research-docs/blob/master/klerosfees12fev2019.pdf` Hence, we have:

$$E[\text{Attacker payoff}] = \left[1 - (1 - yp)^K\right](-F - S_C) + (1 - yp)^K \cdot x.$$

The attacker has no dominant strategy that she should employ regardless of the strategy of the challengers: indeed, if $y = 0$ the attacker's payoff is $x > 0$ (though potentially arbitrarily small), so the only possible dominant strategy would be to always attack. However, if $y = 1$, the attacker's payoff is

$$1 - (1 - p)^K \cdot (-F - S_C - x) + (1 - p)^K x$$

, which approaches $-F - S_C < 0$ for sufficiently large K.

Note that the challenger only pays gas if

- she correctly challenges an inadequate translation - there is a $p \cdot u$ probability of this, or
- she incorrectly challenges a good translation - there is a $(1 - p) \cdot (1 - u)$ probability of this.

(Whereas for the expected challenger reward we consider the possibility that more than one challenger will flag a translation, and only the challenge that is accepted into a block will lead to a reward, here we pessimistically assume that all challengers who try to challenge pay gas. Indeed it is possible that if they challenge at the same time some of their transactions will be reverted leading to a loss of gas.) Hence the total expected gas that a challenger pays is $g[pu + (1 - p)(1 - u) = g[u(2p - 1) + (1 - p)]$.

Then, similar to the considerations of `https://github.com/kleros/research-docs/blob/master/klerosfees12fev2019.pdf`, the challenger's expected payoff is given by

$$E[\text{Challenger payoff}] =$$

$$\frac{u(nF + mx)}{Ky}(1-(1-py)^K) - \frac{(1 - u)F}{Ky}(1-(1-(1-p)y)^K) - e - g[u(2p-1)+(1-p)].$$

If $u = 0$, the challenger's payoff is given by

$$-\frac{F}{Ky}(1 - (1 - (1 - p)y)^K) - e - g(1 - p) < 0.$$

Hence the only possible dominant strategy for the challenger is to never evaluate cases.

However, if $u = 1$ and $K = 1$, the challenger has a payoff of $(nF + mx)p - e - gp$. This is positive if $nF + mx > \frac{e_0 x}{p} - g$, which particularly holds if $m > \frac{e_0}{p}$ and $nF > g$. Then, under these constraints, there are no pure equilibria, so we must have some mixed equilibrium.

Then in equilibrium,

$$y = \left(\frac{(n+1)F + m \cdot x}{(n+1)F + (1+m) \cdot x} \right)^{\left(\frac{1}{K}\right)},$$

$$u = \frac{\left(K \cdot y \cdot (1-p) \cdot g + K \cdot y \cdot z \cdot x + F \cdot \left(1 - (1 - (1-p) \cdot y)^K \right) \right)}{(n \cdot F + m \cdot x) \cdot \left(1 - (1 - p \cdot y)^K \right) + F \cdot \left(1 - (1 - (1-p) \cdot y)^K \right) - (2p-1) \cdot g \cdot K \cdot y}.$$

Then the percentage of malicious translations that go unchallenged in equilibrium is:

$$u(1-yp)^K$$

. We can more explicitly compute

$$u = \frac{\left(K \cdot \frac{1}{p}\left(1 - \left(\frac{(n+1)\cdot F + m\cdot x}{(n+1)\cdot F + (m+1)\cdot x}\right)^{\frac{1}{K}}\right) \cdot (1-p) \cdot g + K \cdot \frac{1}{p}\left(1 - \left(\frac{(n+1)\cdot F + m\cdot x}{(n+1)\cdot F + (m+1)\cdot x}\right)^{\frac{1}{K}}\right) \cdot z \cdot x + F \cdot \left(1 - \left(1 - (1-p) \cdot \frac{1}{p}\left(1 - \left(\frac{(n+1)\cdot F + m\cdot x}{(n+1)\cdot F + (m+1)\cdot x}\right)^{\frac{1}{K}}\right)\right)^{K}\right)\right)}{(n \cdot F + m \cdot x) \cdot \left(1 - \left(1 - p \cdot \frac{1}{p}\left(1 - \left(\frac{(n+1)\cdot F + m\cdot x}{(n+1)\cdot F + (m+1)\cdot x}\right)^{\frac{1}{K}}\right)\right)^{K}\right) + F \cdot \left(1 - \left(1 - (1-p) \cdot \frac{1}{p}\left(1 - \left(\frac{(n+1)\cdot F + m\cdot x}{(n+1)\cdot F + (m+1)\cdot x}\right)^{\frac{1}{K}}\right)\right)^{K}\right) - (2p-1) \cdot g \cdot K \cdot \frac{1}{p}\left(1 - \left(\frac{(n+1)\cdot F + m\cdot x}{(n+1)\cdot F + (m+1)\cdot x}\right)^{\frac{1}{K}}\right)} \tag{1}$$

and

$$(1-yp)^K = \cdot \frac{(n+1) \cdot F + m \cdot x}{(n+1) \cdot F + x\,(1+m)}.$$

Then, one wants to choose values of n and m so that $u(1-yp)^K$ is acceptably low over the various (realistic) values that x might take, both for small scale and larger scale translations.