Some of the recent discussions about integrating the different components of anoma made me think of approaches that enable us to do incremental, but (coarsely) unified, analysis of the whole stack.

Layers of a protocol embedding

When designing a protocol, we need to take into account how it is embedded into its context.

Physics

The physics a protocol rests upon describes the relevant constraints a model of reality imposes on the protocol.

In the case of anoma the constraints can be split into (at least) two classes:

1. Complexity constraints. This contains e.g. intractability assumptions for certain problems, which are relevant for the cryptographic primitives we use.
2. Network constraints. This contains, e.g. the CAP theorem, consensus assumptions and models, the best approximation of time being a vector clock and messages being the only observable objects.

Protocol

The protocol we define is a set of behavioral commitments, i.e. how agents that implement the protocol react to messages of specific form and content and when to send them. By extension, agents expect other agents that implement to protocol to behave the same way.

For these expectations to be fulfilled individual commitments need to beverifiable within the protocol, usually by being bound to its assumed physics in reliable ways.

For example: Cryptographic primitives like signatures, hashes and ZK proofs enable us to verify specific classes of statements, which we can use to construct higher order semantic commitments.

It is important to note that behaving according to the protocol should be (as close as possible to) incentive compatible for agents, if we want to be able to expect adherence.

Out-of-band

Any behavior of agents that falls outside of the behavioral commitments specified in the protocol.

There might be commitments that describe behaviors at the boundary, e.g. how an information oracle ought to function: To use information from outside of the protocol, an oracle assures the validity of some information by wrapping it in a signed message, also containing the oracles identity. This way the oracle stakes its reputation on the validity of its claim.

The protocol could also define (or leave implicit) to fall back to out-of-band governance e.g. to recover from certain failure modes, or for choosing which mechanisms to protocolize.

Protocols as composed games

One approach to model the above could becompositional game theory.

It uses an open game formulation (see Section III of the paper) which admits composition of games, and more importantly the analysis of their properties under parallel and sequential composition.

The framework should also enable us to decompose the larger games we have identified as we learn more.

This might sound similar touniversal composability, with weaker guarantees and more flexibility. This seems helpful to guide our exploration of the design space and help with proposing candidates for an ideal functionality of anoma. Once we have a well defined idealized functionality for anoma, trying to come up with a full UC proof for our construction would still be desirable.

Binding between games and across layers

Looking back at the whole stack, physics, protocol, out-of-band, with the new perspective of trying to decompose it into smaller games, let us revisit a concept which we are going to call binding

.

Binding determines the strength of coupling between games, essentially modelling constraints for theircomposition operator.

In the composed game setting, we should be able to model the composition operator between two games as a binding

game, or move it into the games being composed.

Games that are strictly bound from out-of-band to the protocol should be treated as implicitly in-band. (For physics this is technically true as well, but as long as there is neglegible feedback from the protocol to its physics, we can ignore this.)

Examples:

Cryptography

A cryptographic primitive can be modeled as a game. Assuming its assumptions hold and the implementation is correct, it strongly binds a game composed with it to the physics.

Distributed systems

TODO

Credible commitment devices

Let us first import the definition of acredible mechanism briefly. From the abstract:

“Suppose the auctioneer” (or mechanism designer) “can deviate from the rules provided that no single bidder detects the deviation. A mechanism is credible if it is incentive-compatible for the auctioneer to follow the rules.”

We might be able to introduce a relaxation on this notion of credibility: Assume two mechanisms A

and B

, composed as A \circ B

. A

is a pre-existing mechanism, and a designer wants to establish mechanism B

. Assume all potential participants for B

are already participating in A

.

Then approximate directional credibility

cred(A, B)

of B

in respect to A

can be recovered if the designer of mechanism B

must account for the participants ability to detect deviation, and use this information for enforcement in mechanism A

(see slow games). I don’t expect this to be symmetric, i.e. cred(B, A)

probably has to be analysed sperately.

A credible commitment device (a protocol can be one) can be used to bind some types of games strongly to each other. An approximate commitment device will fall somewhere on the weak-strong scale, proportional to the accuracy of approximation.

Commitment devices can be used to structure the incentives between in-band and out-of-band games.

Correspondence of representations

Let us introduce proofs-of-translation: They prove correspondence between different representations of the same object, e.g. source code and target code, in-memory structs and wire formats.

For any language pair with deterministic translation (e.g. determistic compilation, provably correct parsing), everyone who has access to source and target representations can verify their correspondence.

This can bind games requiring different representations.

What does that mean for us?

Let us derive some (draft) principles for protocol design and implementation:

- The specification needs to construct a hierarchy of games that bind to each other and fundamentally to the physics.
- Compatible implementations are isomorphic up to implementing the behavioral commitments the protocol specifies.
- Behavioral commitments should be defined at the lowest layer where the relevant information to enforce the commitment is available, to maximize generality.

Failure modes of binding and band separation

As designers we try to come up with abstractions that ideally are not leaky, but when dealing with complex, composed systems, mistakes might happen. Mistakes that could happen with the binding between games might look like this:

Behavior that is supposed to be out-of-band and not bound, but which is strongly bound

- The definition of the protocol did not assume the possibility of some external game that, after its discovery makes it incentive compatible for agents to defect out-of-band by participating in it.

Behavior that is supposed to be in-band, but which is not bound

- The credibility of a mechanism relied on some observations being possible, but a new side channel was discovered which enables effective out-of-band coordination, reducing observability.
- Some cryptographic primitive, or its implementation is found to be insecure or broken, removing the binding to physics and invalidating any mechanism relying on the primitive.

Open questions

- Given a specific physics, is there a unique protocol (up to which isomorphism?), which enables us to utilize all optional expressivity induced by the physics, but not more? This should would be a protocol that does not introduce artificial constraints or relies on “unnatural” assumptions which can not be derived from the physics.
- Do separation games exist, or can there only the absence of specific binding games? Could information flow control be related?
- All of the technical details.