Using Role Based Modeling Language to Determine Safety from Advanced Persistent Threats

Andrew Johnson

What are Advanced Persistent Threats? (APTs)

- Cyber threat groups which invest in long-term compromise of their target
- Target organizations, institutions, and governments
- Utilize a wide suite of malware to accomplish mission
- Typically stay hidden within the target's system for a long duration
- Usually allow for future compromise upon completing their mission

APT38

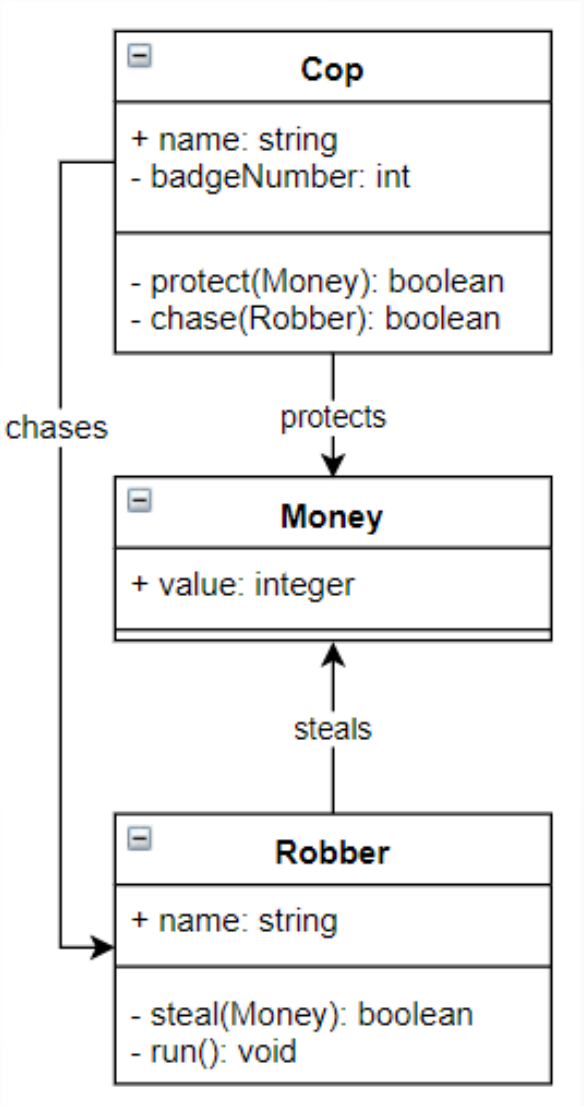
- FireEye is a cyber security company that has identified many APTs and this is the 38th
- APT38 is a North Korean APT, most likely government backed
- Targets financial institutions
- APT38 has attempted the theft of over $1.1 billion since they began their heists in 2015
- APT38 targets SWIFT messaging system for inter-bank communications
- APT38 steals money through fraudulent transactions
- APT38 is willing to become destructive upon being found in a system

Our Goal

- Create algorithm which determines if a security design pattern is present within software
  - A design pattern is a blueprint for a software structure which solves a design problem
- Utilize RBML to model a security design pattern
- Define an algorithm to determine how close a system is to containing this RBML model, and therefore being more secure

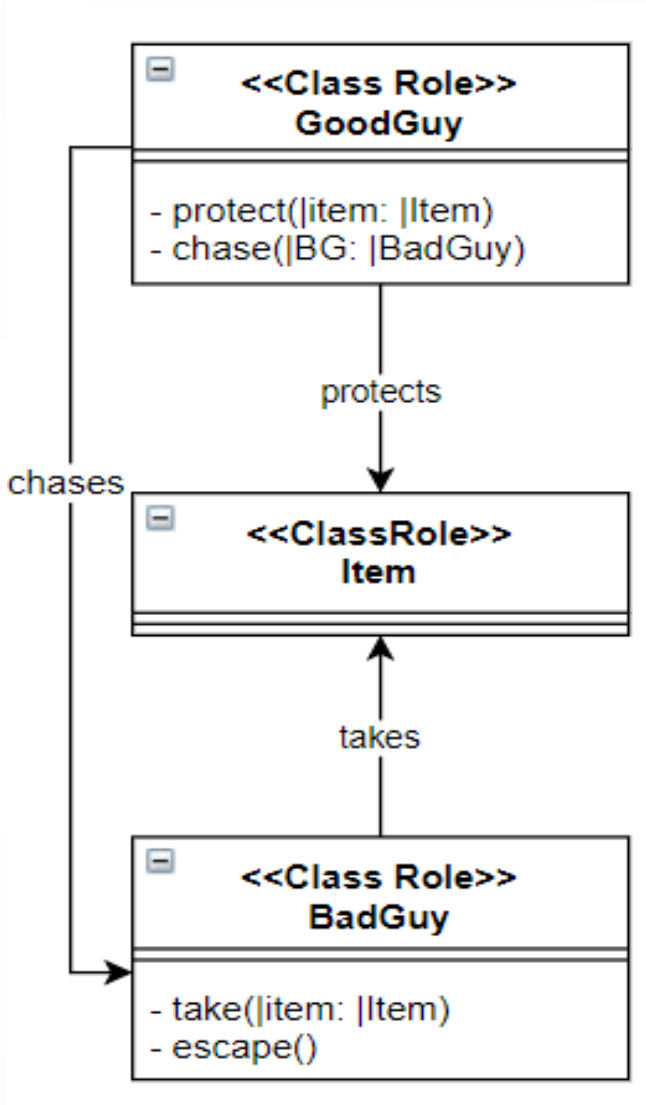
Unified Modeling Language (UML)

- Used to define structure for software
- Allows us to specify structure through a diagram

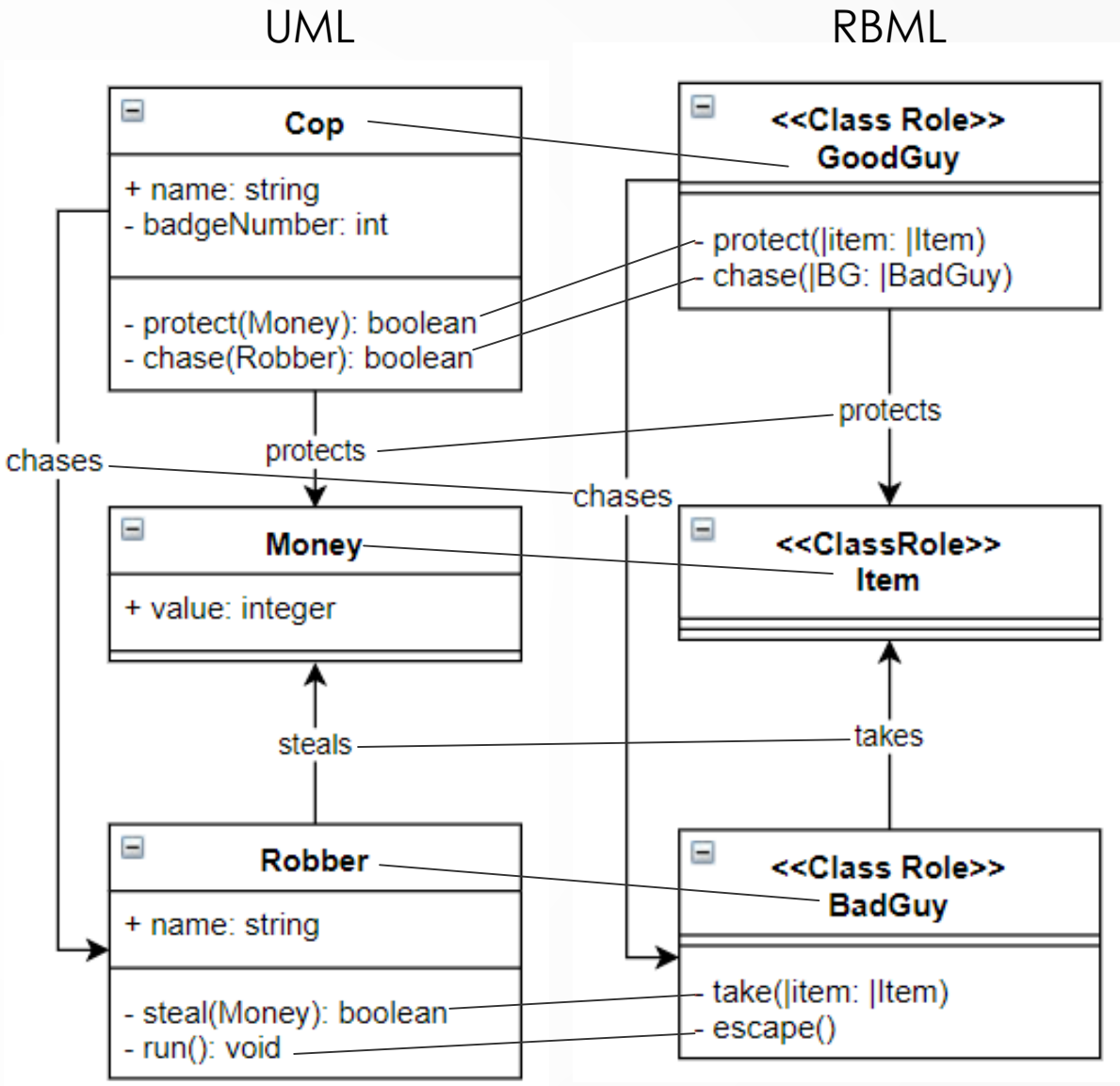


Role Based Modeling Language (RBML)

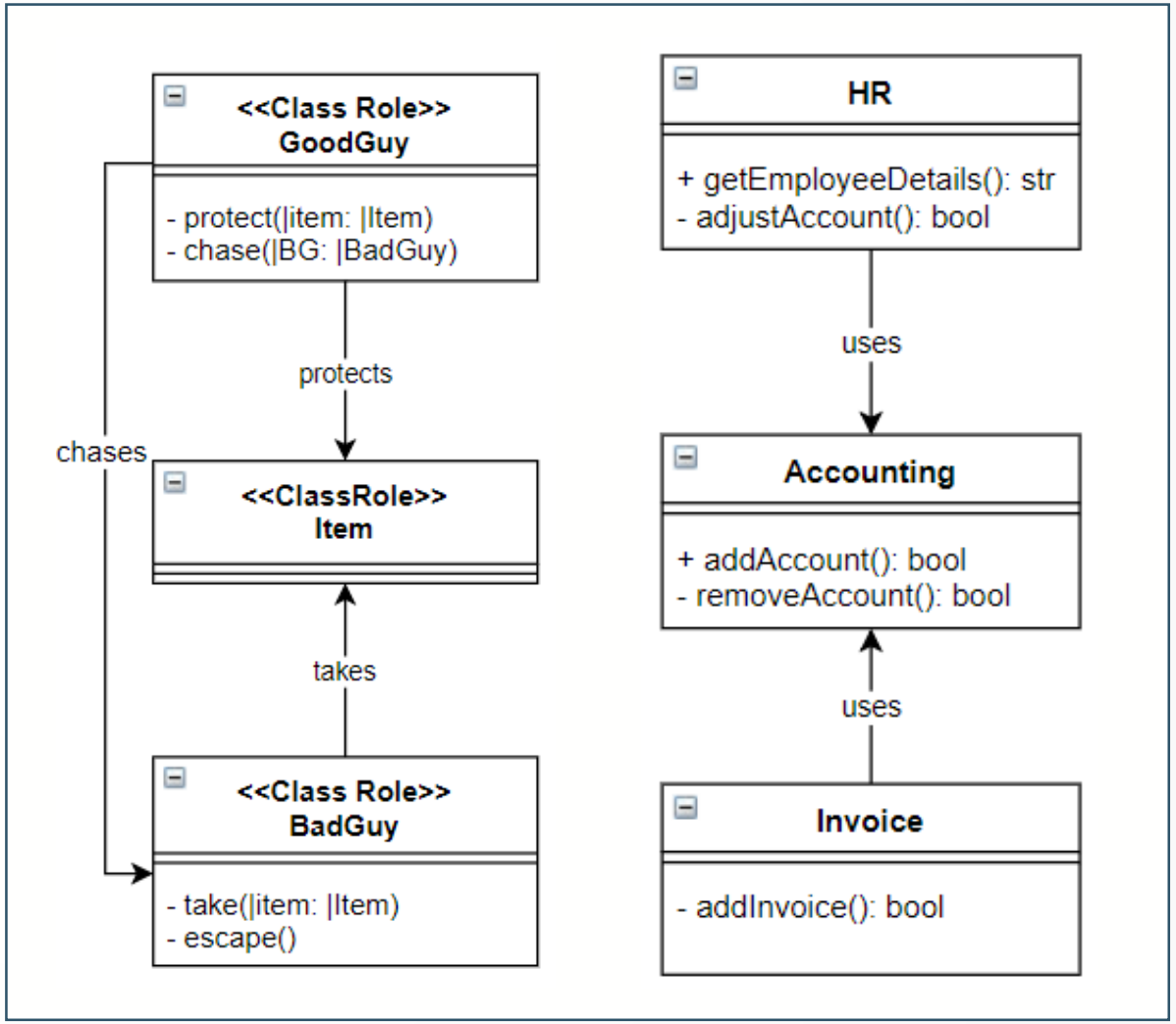
- Defines sets of UML that are valid realizations of a design pattern
- An item is called a role
- Each role has requirements for behavior, relations, and attributes



RBML Example with UML



Example of non-fit UML

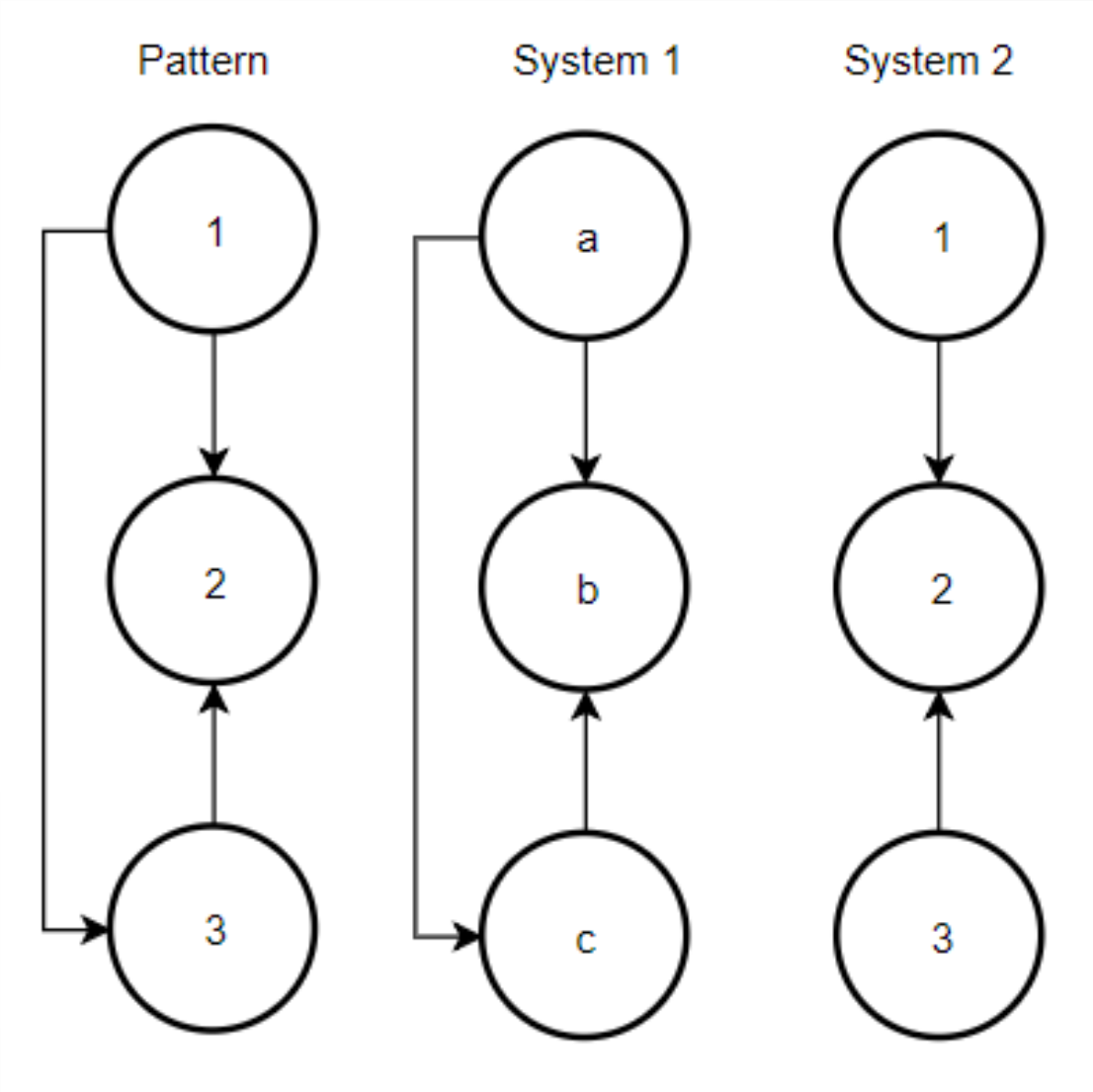


The Johnson Measure

- The Johnson measure is a mathematical metric which gives distance between two graphs.
- If the graphs are isomorphic, or "the same", then the distance is 0
- Let's apply it using our three models

UML/RBML Diagrams as Graphs

- Represent models with graphs where associations are directed edges between two vertices
- Then, represent graphs as adjacency matrices
- System 1 is our Cops and Robbers UML example
- System 2 is the HR/Accounting UML example



Adjacency Matrices

- Pattern : $\begin{matrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{matrix}$
- System 2: $\begin{matrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{matrix}$

- Find the permutation of rows and columns such that the elementwise difference between the matrices is minimized
- In this case, these are the optimal permutations
- We then take the square root of the sum of elements squared (Frobenius Norm), and that is our output

Output and Formal Definition of Johnson Measure

- The Johnson Measure between our pattern and system 2 is 1. That is to say that D(Pattern,System2) = 1.
- Formal definition of the Johnson Measure:
- $D(A,B) = min_p(|P^T AP - B|_f)$
- A, B are the two models being compared
- $min_P()$ indicates we want the *P* such that the statement in parenthesis is minimized
- P, P^T are the permutation matrices
- $|\cdot|_f$ is the Frobenius Norm

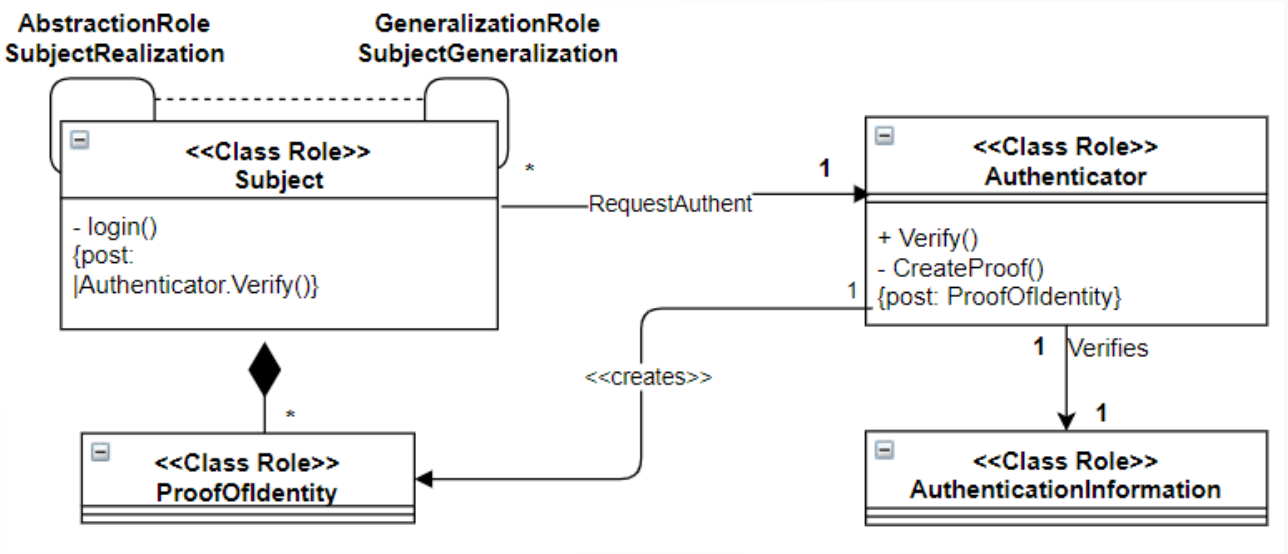
How it Pieces Together

- We may help to assess a system's security with the following method:

- Identify useful security pattern
- Utilize Johnson Measure to identify possible instances of pattern in system
- Compare RBML to possible UML instances to determine if pattern is present

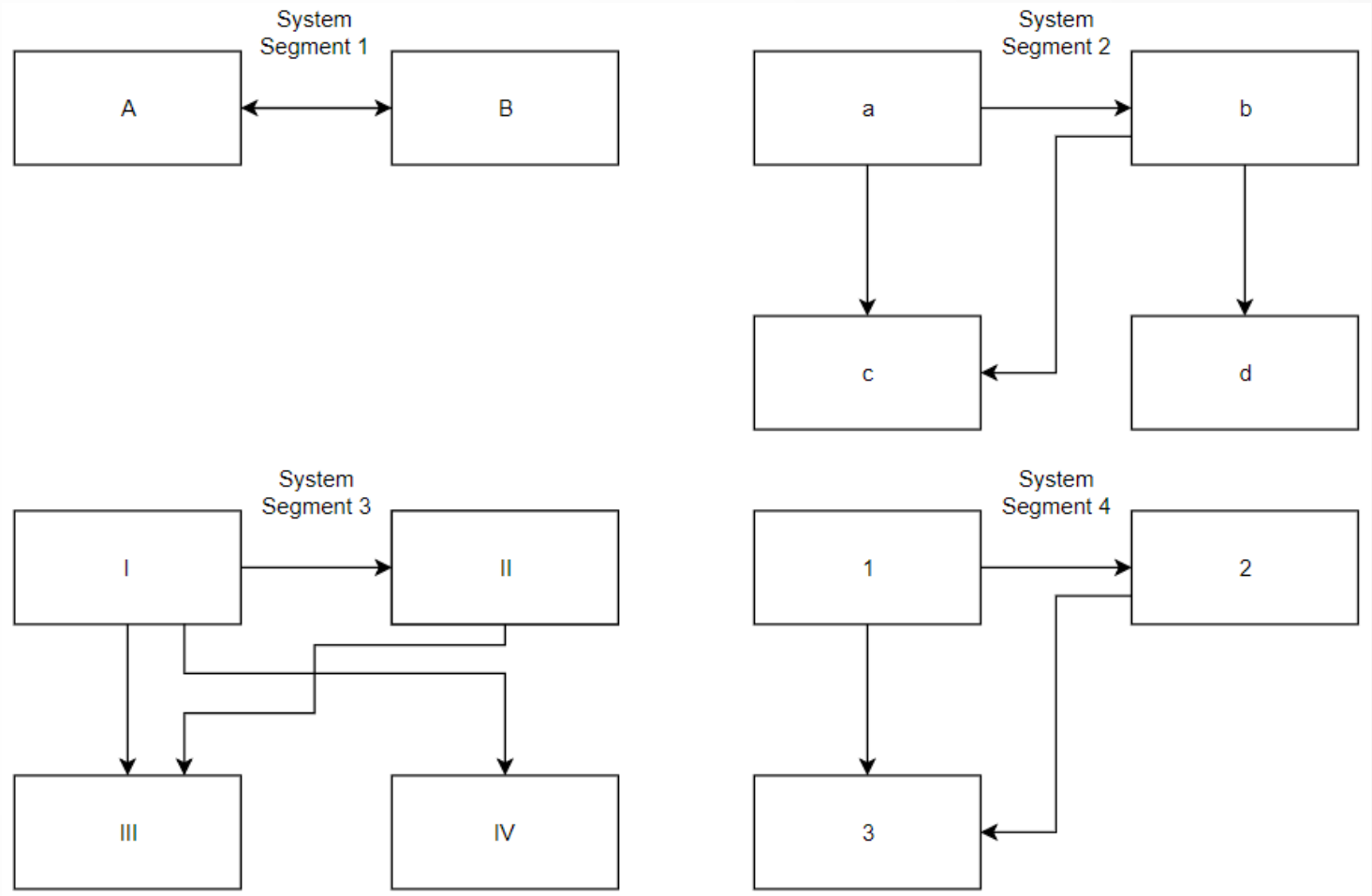
The Authentication Pattern

- Authentication is the verification of identity.
- The premise of this pattern:
  - subject wants to access something
  - Subject provides some authentication information to an authenticator
  - authenticator verifies authentication information
  - Authenticator creates identity for the subject.
- Authentication will help against APT38 because it will stop threat actors from impersonating users



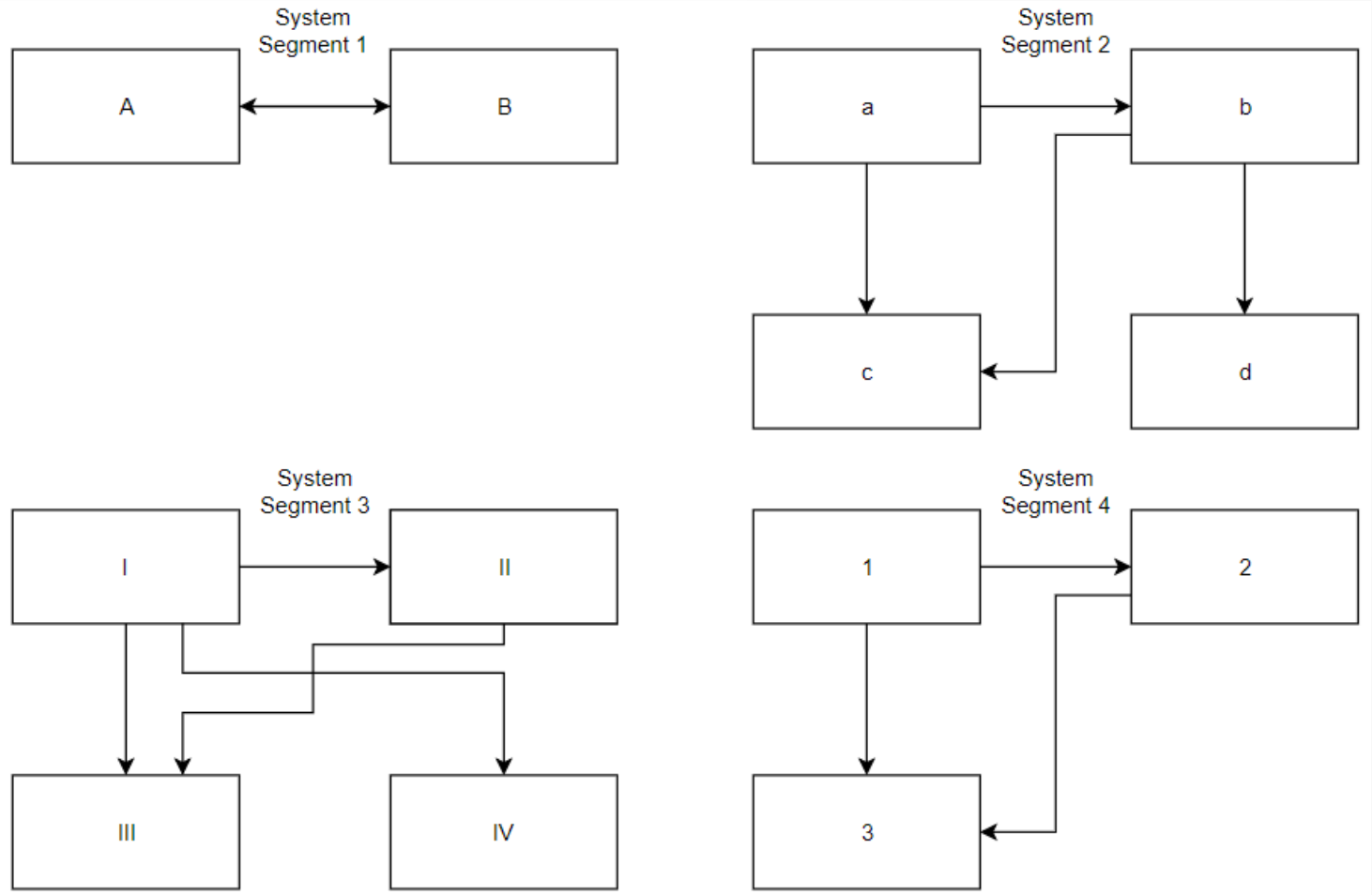
Applying The Johnson Measure Using Authentication Pattern

- We have created several example system segments
- We expect certain output for these segments



Output for System Segments

- Let *P* be the authentication pattern, and S_n refer to the n^{th} system segment.
- $D(P, S_1) = 2$
- $D(P, S_2) = 0$
- $D(P, S_3) = 1.41$
- $D(P, S_4) = 1$



Questions?