CSC 6585

Midterm Exam 2024 - Answers and Rubric

Secure Software Development, Computer Science, Tennessee Tech University

Instructions

This is an open-book, open-notes exam. Answer questions to the best of your ability, working **on your own**. Each question is worth 10 percent of the exam grade. Some questions have multiple parts; all parts have the same weight. You may answer as many questions as you wish, but **only the 10 best answers** will be counted.

Handwritten and scanned or photographed documents are fine, but you are expected to write legibly and to use reasonably correct grammar and spelling; answers that cannot be read or evaluated will not receive any points. Please do not create ASCII art diagrams.

Be sure to state your assumptions explicitly.

Max: 100 / Avg: 97.0

Maximum and average score is given for each question to two significant digits. The average includes zeros where a question was not attempted. The questions are ranked by average, and the rank is given, as well.

Quick Links: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Questions

1. Your company uses an online data-sharing system for communicating in real-time with partners. The online system is pre-configured, provisioned, and maintained by a third party.

   Is the data-sharing system inside or outside your company's system boundary? Justify your answer.

   Max: 10. / Avg: 9.5 / Rank: 4

   See Motivation.

   You could give either answer, provided you supported your answer adequately.

   The data-sharing system is outside the boundary. You do not have direct control over the data-sharing system so you must treat it as outside your system boundary and specify the security requirements of the flows to and from the data-sharing system. Remember that the boundary is defined by the interfaces, so the interfaces with the data-sharing system define the boundary here.

   The data-sharing system is inside the boundary. There are probably lots of parts of your system that you recognize as such but do not directly maintain, like third-party libraries, monitoring tools, etc. Because it is integral to the functioning of your enterprise, you must treat it as inside the system boundary and take responsibility for auditing its security properties through contract enforcement with the third-party supplier.

The purpose of this question is to demonstrate your understanding of the system boundary and also to point out the (somewhat) arbitrary nature of the boundary. Everything is connected, and where you choose to draw the boundary defines your security posture, needs, and policies.

It needs to be entirely in or entirely out of the system. A "partial in/out" answer is not good because it doesn't make a clear decision and runs the risk of treating the security incompletely or inconsistently. (In for X, but out for Y…)

You got half credit for an answer that (1) specifies whether it is in or out and half credit for (2) justifying the answer.

2. You work for a logistics company and maintain your corporate database on a cloud provider's platform. Consider the following.

    - Serrice Technology provides telematics (telematics is hardware and software for monitoring a fleet of vehicles) and pushes information to your cloud database. This information includes vehicle location, weight, and health.

    - Ascension Financial Services has a data feed from your cloud database. This information includes contract fulfillment, payment processing, and accounts due.

    - You communicate with your drivers using a third-party communication provider, Sonax Industries.

    Draw the components of this system, identify trust boundaries, and label all flows. Be sure to state your assumptions.

    Max: 10. / Avg: 6.5 / Rank: 13

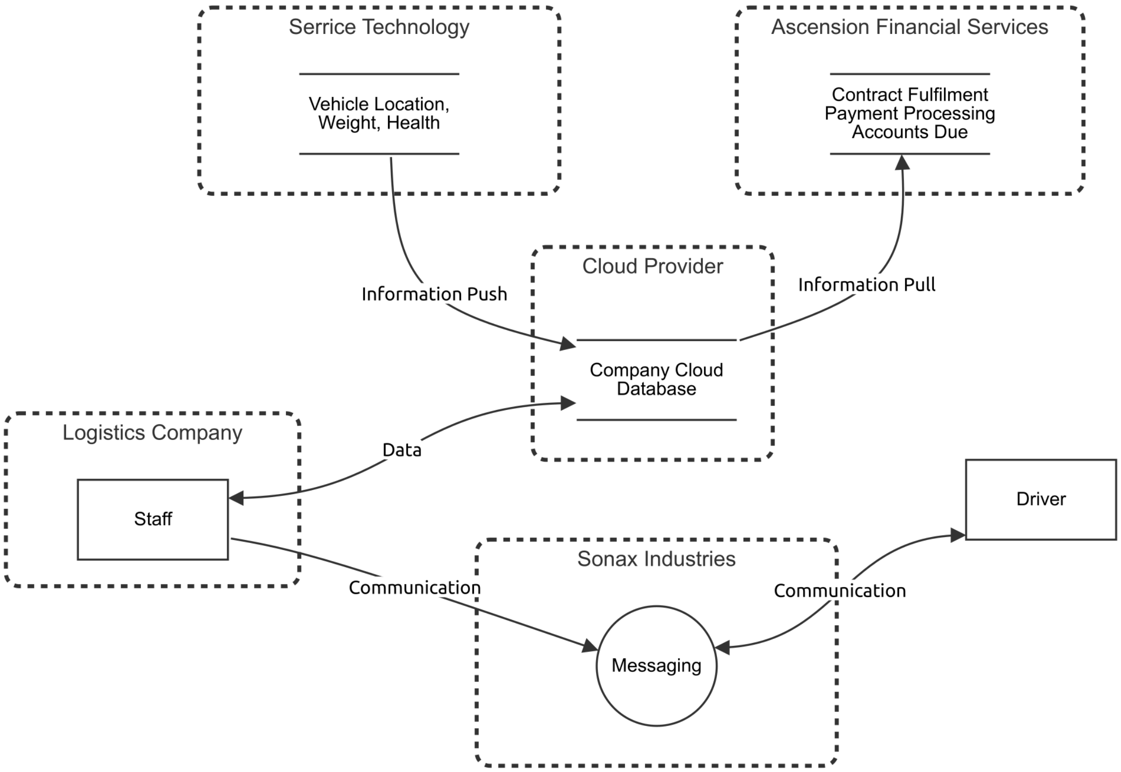
    See Models 1.

    You could use the threat modeling tool to do this, or you could create a diagram in another program. Your diagram needed to include the following items to get full credit.

    - Serrice Technology's telematics, with boundary
    - Ascension Financial Systems accounting, with boundary
    - Sonax Industries communications, with boundary
    - The cloud provider, with boundary
    - Your drivers, with or without boundary
    - The logistics company itself, with boundary

    So long as a flow crossed a trust boundary whenever it needed to, you did not have to draw boundaries around every item above. You can identify boundaries with lines instead of boxes, but you need to show that the trust boundary between, say, Serrice Technology and your Cloud Provider is distinct from the trust boundary between Ascension Financial Systems and your cloud provider.

    The result might look like the following.



The purpose of this is to show the relationship of the different trust boundaries.

3. A software system has a large set of configuration options, stored as JSON files with one JSON file per component. These files are stored in a git repository, allowing users to switch configurations (using branches) and roll back to prior configurations if problems are detected.

   Max: 10. / Avg: 8.0 / Rank: 7

   See Basics.

   a. The git repository can be stored locally in the software's configuration folder (an application-specific folder under `AppData` on Windows, `Library/Preferences` on macOS, and `~/.config` on Linux). What, if any, concerns would you have with this arrangement?

      The repository is accessible by the user, any deputy of the user, or anyone with that user's access. If the account itself is not properly secured, then the database may be accessible by anyone. The consequences of this depend on the content. An adversary may change paths in the configuration to insert malicious code, obtain security tokens or credentials stored in the database, or corrupt the database leading to a denial of service.

      More directly, suppose the system is an editor that must specify the location of the chosen interpreter, such as Java or Python. A malicious interpreter can be installed in front of the actual interpreter and used to modify input/output or to leak secrets from the system.

If this option is chosen, then for critical operations (for example, with privilege elevation) you need to treat the database contents as untrusted.

Additional notes. The repository contains a history, so even if sensitive items are removed, they may persist in the history.

To get full credit for this you needed to note the possibility of tampering with the database. The focus is, after all, security.,

b. The git repository can track a remote cloud-hosted repository, enabling settings sync across multiple machines. What, if any, concerns would you have with this arrangement, distinct from those identified in part a?

The additional concerns are that the repository may compromised by a hack of the provider, someone may obtain credentials allowing them to tamper with the repository remotely.

Since the repository is cloud-hosted and shared among multiple machines, tampering with the repository allows the damage to spread to all machines that share the database.

To get full credit for this you needed to note the possibility of remote tampering with this database.

The purpose of this question is to get you to think about the security of data based on where it is stored ("at rest") and of configuration information more generally.

4. Consider this proposed definition. "A software system is secure iff there is no series of inputs to the system that results in output of the content of a file unless the final input in the series originates from the file owner."

Now consider the CIA triad. For each of the elements of the triad, does the above definition address that element? Explain.

Max: 10. / Avg: 8.8 / Rank: 6

See Models 1.

Confidentiality. This definition addresses confidentiality by requiring that the data owner take the final action on the path to data disclosure.

Integrity. This definition doesn't address integrity since it says nothing about the data itself.

Availability. This definition doesn't address availability since a system that outputs nothing ever meets the definition.

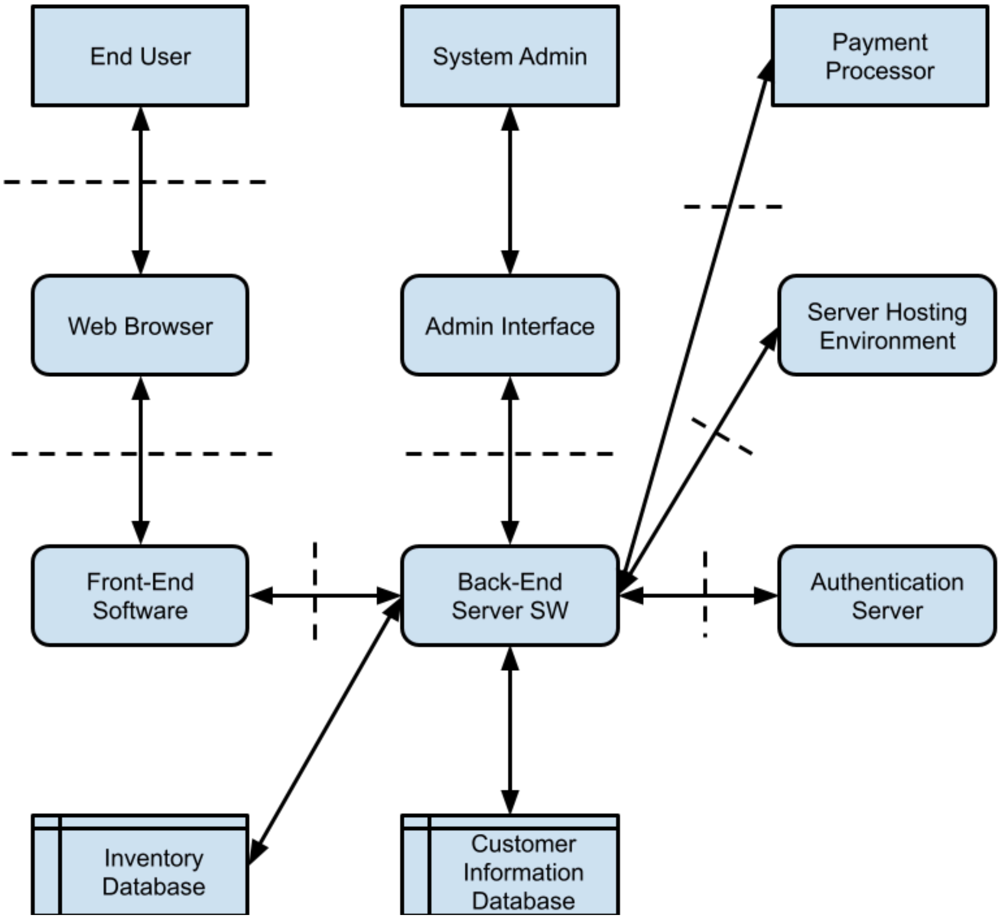
The above is all that is necessary for full credit.

The purpose of this question is to consider the difficulty of defining security. Note that nothing is said about the action taken by the owner. What if the path leads to a state where the final action is simply the owner logging in (using, perhaps, the shell as a deputy to leak secrets by modifying the shell's startup scripts)? It should be clear that this definition is insufficient, even for confidentiality.

5. Consider a simple web store. A back-end server provides the business logic, while the front-end runs in a web browser.

Max: 10. / Avg: 5.3 / Rank: 14

a. Draw a block diagram of this scenario, labeling all information flows. You should include the end user, web browser, front-end software, back-end server software, server hosting environment, inventory database, authentication server, customer information database, payment processing contractor, the administrator interface (for simplicity we will assume it runs natively on the server), and the system administrator.



The main thing you needed to do was to show the components and the flows.

b. Specify the system boundary for the server software by identifying its interfaces.

i. Interface with the customer information database

ii. Interface with the inventory database

iii. Interface with the authentication server

iv. Interface with the front-end software

v. Interface with the administrative interface

c. For every control or information flow in your diagram for part a, determine if the flow crosses a trust boundary. Consider bidirectional flows to be a single flow for the purpose of this part.

Trust boundaries are shown in the diagram with dashed lines.

6. A piece of financial software has the following logic. Here `RESULT` is a result code, `Acct` is an account structure, and the `LOCK` and `UNLOCK` macros prevent a race condition. Note that `static` functions in C are private to the containing module.

Max: 10. / Avg: 9.1 / Rank: 5

```
static RESULT _deduct(Acct * account, int value)
{
    if (account->balance < value) {
        return INSUFFICIENT_FUNDS;
    }
    account->balance -= value;
    return OK;
}

RESULT deduct(Acct * account, int value)
{
    LOCK(account);
    RESULT result = _deduct(account, value);
    UNLOCK(account);
    return result;
}
```

a. What happens when the value is a negative number and the account balance is positive?

First, we check that the balance is less than the value. However, since the balance is positive and the value is negative, this will be false and the body of the if-statement will be skipped.

Next, the value will be subtracted from the account balance. In this case, because the value is negative, its absolute value will be added to the account balance.

It is also possible, since we are using fixed-width computer integers, that the result is an overflow or underflow, resulting in a negative balance. For instance, suppose the account balance is $10 (represented as 1000 using fixed-precision arithmetic) and we "deduct" a negative value close to the limit for the type. Then the result can be negative due to underflow.

To get credit for this question you needed to note that the absolute value will be added to the account balance.

b. How would you recommend addressing this problem in the given code? Be specific about how you would change the code.

There are many possibilities to address this.

You could test for and reject a negative value in either the deduct or _deduct function. Here we assume BAD_VALUE is defined somewhere.

```
if (value < 0) {
    return BAD_VALUE;
}
```

You could add an assertion to the program to catch this condition. This might not be what you want, since this software needs to keep running. If you did add an assertion, you also need to add a handler to catch the assertion at a higher level. The same is true for using an exception in C++.

You could replace the type for value with an unsigned integer. This would allow the compiler to catch cases where potential negative values are being passed, but may require more changes in the API.

You needed to answer with one of the above to get full credit, and needed to be explicit about what to do, since the question asked for that. Simply "check the value" is not sufficient for credit.

The purpose of this question was to get you thinking about integer security.

7. Canada and the US decide to connect information processing systems based on the Bell-LaPadula model. Fortunately, the two countries have exactly the same sensitivity levels. To preserve state secrets, both countries introduce an additional security modifier: CAN-ONLY for Canada and US-ONLY for the US, and corresponding clearance tags: CAN for Canada and US for the US. These sensitivity levels are independent of the usual TOP SECRET, SECRET, CONFIDENTIAL, UNCLASSIFIED forms.

Is the resulting system still Bell-LaPadula secure? Consider each of the required properties and whether it holds for the composite system.

Max: 10. / Avg: 7.5 / Rank: 11

See Models 1.

The Bell-LaPadula system has three properties that have to be satisfied: the simple security property (no read up), the star property (no write down), and the strong star property (no read/write up or down).

We assume that only Canadian citizens can have CAN, and that only US citizens can have US. We further assume that dual citizens are not permitted to have either CAN or US. Finally, we assume that if an object has both CAN and US this is the same as having neither, and ignore that case. WLOG we will consider only Americans, as dual citizens are a subset of this and the two countries are symmetric. If neither the subject nor the object has the CAN or US modifier, then this is exactly the original Bell-LaPadula model, and all three properties hold. We need only consider the case where the subject, object, or both have a modifier.

Suppose that the subject has access (c_s,k_s) and the object has access (c_o,k_o). Then we have the following cases.

- If $c_s \geq c_o$ and $k_s \supseteq k_o$ then the subject's access dominates the object's access. In this case the subject can "read down" to the object. This satisfies the requirements for the simple security property.
- If $c_o \geq c_s$ and $k_o \supseteq k_s$ then the object's access dominates the subject's access. In this case the subject can "write up" to the object. This satisfies the requirements for the star property.
- If neither $k_s \supseteq k_o$ or $k_o \supseteq k_s$, then neither the object nor the subject dominates, and no access is permitted.

From these it is also clear that the strong star property also holds, and the new system is Bell-LaPadula secure.

It was also sufficient to note that the modifiers fit in the set K of the model, and thus really changed nothing about the model.

You just needed to demonstrate that you understood the Bell-LaPadula access controls to get credit for this question.

8. Recall the five functions for physical security and classify each of the following security measures to those functions. Identify which function or functions are addressed by each.

   Max: 10. / Avg: 9.8 / Rank: 3

   See Security 1.

   a. A login screen states that the system is for official use only by US government employees, and that misuse of the system could result in fines or jail time.

      Deter: prevent an action by a credible threat

      You needed to answer deter to get credit.

   b. After three successive login failures the system logs you out and blocks your IP address.

      Detect: (possibly) the announcement of a potentially malicious act

      Respond: respond with the force necessary to stop the advancement of the adversary

      Neutralize: render the adversary incapable of interfering with operations

      You needed to answer either (or both) of respond and neutralize to get credit. Detect is a weak answer because it is not stated whether or not anyone is notified. Delay is similarly a weak answer.

   c. After two login failures, the system forces you to wait a full minute before trying again.

      Delay: impede the adversary from completing a malicious act

You needed to answer delay to get credit.

d. The system detects when a specific file is read and, if that is detected, the stored disk encryption keys are randomized and the system is shut down.

   Detect: (possibly) the announcement of a potentially malicious act

   Respond: respond with the force necessary to stop the advancement of the adversary

   Neutralize: render the adversary incapable of interfering with operations

   You needed to answer either (or both) of respond and neutralize to get credit. There is a strong case for neutralize, but if you missed it that's okay. Detect is a weak answer because it is not stated whether or not anyone is notified.

The purpose was to demonstrate that you understood the five functions and how they can be applied to any system, not just physical systems.

9. Recall the NIST five functions and classify each of the following security measures as one of those functions. Identify which function or functions are addressed by each.

   Max: 10. / Avg: 9.9 / Rank: 2

   See Security 2.

   a. You hire a lawyer to advise you about the relevant laws governing the specific data you intend to collect and store.

      Identify: develop an organizational understanding of managing cyber risk

      You needed to answer identify to get credit.

   b. Your system periodically autosaves its current state to a git repository.

      Protect: limit or contain the impact of cybersecurity events

      You needed to answer protect to get credit. Recovery is not a good answer here because, while this may support recovery, that is not the function of this control.

   c. The system implements a "watchdog timer." If the timer ever counts down to zero, an alarm is reported through the endpoint monitoring software.

      Detect: identify the occurrence of a cybersecurity event in a timely manner

      You needed to answer detect to get credit.

   d. Your IT team runs a "tabletop" exercise with the scenario that you detect an intrusion in the network.

      Respond: maintain plans for resilience and to restore services

      You needed to answer respond to get credit. Protect mentions having a plan, but does not discuss testing that plan and, as such, is a weak answer here.

The purpose of this question was to demonstrate that you understood the NIST five functions and how the apply to systems.

10. A software system checks the host key of a connecting machine and requires that the user identify themselves with both a password and a 2FA token. For as long as the user is logged in, the software periodically sends a challenge to the remote host and then checks the reply to re-authenticate the connecting machine. While the user is interacting with the system, biometrics are collected and compared to the user's baseline.

    Is this an example of defense in depth? Why or why not?

    Max: 10. / Avg: 7.6 / Rank: 10

    See Security 2.

    Defense in depth requires two things: (1) defensive measures exist at different layers of the system and (2) the defensive measures are independent.

    The first criterion is satisfied because we see defensive measures at login, periodically during connection, and at system interaction.

    The second criterion is satisfied because the measures (2FA, host authentication, biometrics) are independent.

    You needed to answer yes and apply the two criteria to get full credit. If you didn't explicitly point out independence, you lost half the points.

    The purpose of this was to demonstrate that you understand defense in depth.

11. Consider the CIA model of Confidentiality, Integrity, and Availability. Different contexts require different security measures.

    Max: 10. / Avg: 10. / Rank: 1

    See Models 1.

    For each of these, you should think about the consequences of failure to meet a criterion.

    a. Give an example of a system where confidentiality is more important than the other two properties. Explain.

       Your answer will vary, but one example is iLearn. Confidentiality is the primary concern here. If the data is incorrect (integrity), then the impact is low and we can address it. If the system is down (availability), then we can record information on paper or in a spreadsheet and enter the data when the system recovers.

       Surprisingly, healthcare systems are another example. It might seem that the consequences of bad data (integrity) are high, but health systems implement redundancies to account for this. The seemingly redundant forms and questions you have to answer protect against bad data in the system.

Good answers here will include systems that must keep their data private to accomplish their mission.

b. Give an example of a system where integrity is more important than the other two properties. Explain.

Most financial systems value integrity over the other properties. Typically there are financial reporting requirements so, while confidentiality is important, it is possible that a third party could leak the information. However, if a bank loses track of account data it is almost certainly the end of that bank.

Good answers here will include systems that focus on quality of information to accomplish their mission.

c. Give an example of a system where availability is more important than the other two properties. Explain.

The power grid is an example of such a system. The grid must continue to deliver power no matter what. If the wrong person is billed (integrity) or information is leaked (confidentiality) there are processes to recover.

Good answers here will be any form of critical infrastructure.

You were judged entirely on your explanation. You needed to justify why the given property was more important than the others to get credit.

12. Consider the HRU command definitions and the safety rule, and let A be the access control matrix.

Max: 10. / Avg: 7.0 / Rank: 12

See Models 3.

a. Create a command foo that accepts two subjects, a and b, and an object named X. The command checks that A(a,b) contains FOO and that A(b,a) does not contain FOO and, if both are true, it adds object X with owner a.

You needed to follow the structure given in slides (or paper), but I was pretty forgiving on this since the paper also used a few shortcuts.

```
command FOO($a$, $b$, $X$)
        if FOO $\in$ A($a$,$b$) and FOO $\notin$ A($b$,$a$)
        then
                Create object $X$
                Enter right owner into A($a$,$X$)
end
```

This answer is careful to use the precise wording found on the slides, but some leeway was granted if I understood what you were indicating. Note that you needed to check for FOO (whatever that token meant) in the table. Remember: the tokens are arbitrary, and chosen to capture the system being modeled.

You needed to get something very close to the above for full credit. Omissions or errors lost points as follows. Basic structure: 1 point. Checking for FOO: 1 point.

Checking for absence of FOO: 1 point. Creating the object: 1 point. Entering the right in the table: 1 point. If you added extra logic or actions, you also lost a point.

b. Consider the command you wrote for part a. Does the command ever leak the right FOO from the system? That is, is it safe with respect to the right FOO?

The command never adds FOO to the table, so yes, it is safe with respect to the right FOO.

You only needed to say "yes" here for full credit. If you explained also, great! If you didn't directly answer the question, you lost a point.

The purpose of this was to demonstrate that you understood what the access matrix was, and how it was used in HRU (and other models).

13. The HRU paper shows that, given a Turing machine, we can convert the question "Does this Turing machine halt on all inputs?" to a protection system and the question of whether it is safe. Thus if we could answer the latter, we would also be able to answer the former, and thus solve the Halting Problem. Since we are confident the Halting Problem cannot be solved in general, it follows that it is undecidable whether a given protection system is safe.

A friend shows you a design for a protection system and tells you that they can prove it is safe. Is that possible? Why or why not?

Max: 10. / Avg: 5.2 / Rank: 15

See Advanced Models 3, which was covered during the exam period, and also discussions of decidability elsewhere in the course.

It is possible. A particular protection system may be proven to be safe, even though no decision procedure can decide safety for all protection systems. As a trivial example, consider a protection system that does not allow the addition of any rights to the table. Such a system is trivially safe.

You needed to (1) say it was possible and (2) explain for full credit.

The purpose of this was to demonstrate that you understood the consequences of the HRU proof.

14. Your goal is to gain access to my iLearn account for this class. Build an attack tree starting with this goal. The leaves of the tree should be specific actions, but you do not need to be more detailed than the example given in class.

Max: 10. / Avg: 7.7 / Rank: 8

See Threats 3.

Your answer will vary, but your attack tree needs to start with "access professor's iLearn account." You needed to demonstrate at least two different paths to get access and end with specific goals on the leaves of the tree.

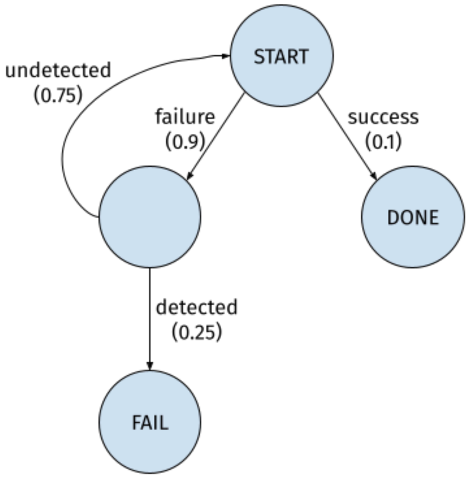
You either got credit for this or not.

15. An attack has a 10% chance of succeeding and, when unsuccessful, there is a 25% chance of being detected. Assume detection results in a change to the system preventing the attack in the future. Create an attack graph that represents this situation and label the edges with actions (success, failure, detection, no detection) and probabilities. Assume the attack will be repeated until it either succeeds or is detected, and capture that in your graph.

    Max: 10. / Avg: 7.7 / Rank: 8

    See Threats 3.

    Your answer will vary, especially with the labels on your graph, but the basic structure must match the following skeleton.



    Of particular importance is the recurrence loop back to the top. This is explicitly called for in the question.

    The structure was worth half the points, and the probabilities were worth the other half. If you failed to label a line, I assumed it was unconditional (probability 1).

16. [Extra 10 points] What is the overall probability that the attack in question 15 eventually succeeds?

    Success: 10

    Consider the probability that the attack succeeds on the first attempt. That has probability 0.1. Now, suppose the first attempt fails, but the second attempt succeeds. This has probability (0.9)(0.75)(0.1), because we fail, are undetected, and then try again and succeed. Note that these two events are independent.

    The probability that we succeed after *n* failures is therefore given by the following equation.

$$S_n = [(0.9)(0.75)]^n \, (0.1)$$

    Since every one of the possibilities is independent, the total probability of success is the sum of the infinite series.

$$S = \sum_{n=0}^{\infty} S_n = \sum_{n=0}^{\infty} \left[(0.9)(0.75)\right]^n (0.1)$$

This is a geometric series and has a well-known solution.

$$S = \frac{0.1}{1 - (0.9)(0.75)} \approx 0.308$$

Overall there is roughly a 31% chance of eventual success.

You needed to get the above answer to get credit. No partial credit was given for this question. If you were successful, the 10 points were added after the exam was otherwise graded, with the total points capped at 100.

The purpose of this was to illustrate the impact of iteration on the probability of an attack, and thus the asymetric nature of attacks. The adversary can try repeatedly and only has to succeed once, and even a low-probability attack may be a threat.

You can play with the values of success and detection. Even if detection is raised to 50%, the adversary still has nearly a 20% chance of success. If the attack and detection both have a 50% chance, then the adversary is nearly guaranteed to succeed (>0.9).