# On $g$-vulnerability

Mário S. Alvim
(msalvim@dcc.ufmg.br)

Quantitative Information Flow

DCC-UFMG
(2021/1)

# Introduction

- We have seen Bayes vulnerability measures one basic threat to a secret $X$:

  The risk that the adversary could correctly guess it in one try.

- But there are many other operational scenarios that we could consider.

  For instance, the adversary might benefit from:

  (a) guessing only part of $X$,

  (b) an approximate value of $X$,

  (c) a property of $X$, or

  (d) from guessing $X$ correctly within a fixed number of tries.

  Also she might be penalized for making an incorrect guess.

- To accommodate this multiplicity of possible operational scenarios, here we cover a decision-theoretic framework of $g$-vulnerability.

# Basic definitions

# Basic definitions

- The perspective of $g$-vulnerability is:

  **Knowledge about a secret $X$ is important only to the extent that it can be exploited by an adversary, enabling her to take some action that rewards her.**

- For instance:

  (a) knowing a user's password enables an adversary to log in as that user;

  (b) knowing a combination enables her to open a safe; and

  (c) knowing a sentry's location enables her to enter a facility without being detected.

- We model the adversary's operational scenario with a nonempty set $\mathcal{W}$ of **actions** that she could take.

  Sometimes (but not always) actions may be **guesses** about the secret, i.e., $\mathcal{W}{=}\mathcal{X}$.

# Basic definitions

- To complete the model of the operational scenario, we specify a notion of reward.

## Definition (3.1)

Given a finite, nonempty set $\mathcal{X}$ (of possible secret values) and a nonempty set $\mathcal{W}$ (of possible actions), a **gain function** is a function $g\colon \mathcal{W}\times\mathcal{X} \to \mathbb{R}$.

- The value $g(w, x)$ specifies the gain that the adversary achieves by:
  - taking action $w$,
  - when the value of the secret is $x$.
- It is common to limit the range of $g$ to be the interval $[0, 1]$, s.t.
  - $g(w, x) = 0$ when $w$ is a "worthless" action, and
  - $g(w, x) = 1$ when $w$ is a "best" action.

  However, in general, it makes sense to allow arbitrary gain values.

# Basic definitions

- The actions $\mathcal{W}$ correspond to the actual adversarial actions possible:
  (a) "Enter 12913 on the keypad."
  (b) "Try to enter the fort through the south gate."
- A gain function $g: \mathcal{W} \times \mathcal{X} \to \mathbb{R}$ abstracts from the details of what the actions are, as that doesn't affect the quantification of threat.

  Still, a gain function $g$ is relevant to a particular operational scenario only if:
  1. there really are actual actions corresponding to the elements of $\mathcal{W}$; and
  2. whose effectiveness is correctly modeled by $g$.
- When $\mathcal{W}$ is finite, we can represent a gain function $g$ as a matrix G indexed by $\mathcal{W} \times \mathcal{X}$ so that

$$\mathsf{G}_{w,x} = g(w, x)$$

and

  - the rows correspond to actions; and
  - the columns correspond to secrets.

# Basic definitions

- Now we are ready to define $g$-vulnerability.

## Definition (3.2)

The $g$-**vulnerability** of $\pi$ is defined as

$$V_g(\pi) \ := \ \max_{w \in \mathcal{W}} \sum_{x \in \mathcal{X}} \pi_x \, g(w, x) \quad .$$

If $\mathcal{W}$ is infinite then the max should be replaced by sup.

The key idea is that when $X$ has distribution $\pi$, a smart adversary should choose an action $w$ that maximizes her expected gain

$$\sum_{x \in \mathcal{X}} \pi_x \, g(w, x)$$

with respect to $\pi$.

# Basic definitions

- Example 1 Consider a scenario having:
  - secret set $\mathcal{X} = \{x_1, x_2\}$;
  - prior $\pi = (0.3, 0.7)$;
  - action set $\mathcal{W} = \{w_1, w_2, w_3, w_4, w_5\}$; and
  - gain function $g: \mathcal{W} \times \mathcal{X} \to \mathbb{R}$ having values as in

| G | $x_1$ | $x_2$ |
|---|---|---|
| $w_1$ | $-1.0$ | 1.0 |
| $w_2$ | 0.0 | 0.5 |
| $w_3$ | 0.4 | 0.1 |
| $w_4$ | 0.8 | $-0.9$ |
| $w_5$ | 0.1 | 0.2 |

.

Compute the value of $V_g(\pi)$.

- Example 1 (Continued)

**Solution.** To compute the value of $V_g(\pi)$, we start by computing the expected gain for each possible action $w$ in $\mathcal{W}$, given by, for each one,

$$\sum_{x \in \mathcal{X}} \pi_x \, g(w, x) \ .$$

The results are as follows.

$$\begin{aligned}
\pi_{x_1} g(w_1, x_1) + \pi_{x_2} g(w_1, x_2) &= 0.3 \cdot (-1.0) + 0.7 \cdot 1.0 &&= 0.40 \\
\pi_{x_1} g(w_2, x_1) + \pi_{x_2} g(w_2, x_2) &= 0.3 \cdot 0.0 + 0.7 \cdot 0.5 &&= 0.35 \\
\pi_{x_1} g(w_3, x_1) + \pi_{x_2} g(w_3, x_2) &= 0.3 \cdot 0.4 + 0.7 \cdot 0.1 &&= 0.19 \\
\pi_{x_1} g(w_4, x_1) + \pi_{x_2} g(w_4, x_2) &= 0.3 \cdot 0.8 + 0.7 \cdot (-0.9) &&= -0.39 \\
\pi_{x_1} g(w_5, x_1) + \pi_{x_2} g(w_5, x_2) &= 0.3 \cdot 0.1 + 0.7 \cdot 0.2 &&= 0.17
\end{aligned}$$

Thus we find that $w_1$ is the best action and $V_g(\pi) = 0.4$ . ◁

# Basic definitions

- Sometimes the adversary "loses" by performing action $w$ when the secret is $x$.

  The smaller $g(w, x)$ is (including negative values), the higher the loss.

- When allowing negative values for $g(w, x)$, we still require that the adversary's maximum expected gain $V_g$ should always be non-negative, so that 0 represents the best case of "no vulnerability".

  In this course we restrict our attention to the class of gain functions $g$ such that $V_g$ is always non-negative, denoted by $\mathbb{G}\mathcal{X}$.

# Basic definitions

- Another useful representation of a gain function is viewing each action as a vector

$$w \in \mathbb{R}^{|\mathcal{X}|} ,$$

whose elements are thecorresponding gains, that is

$$w_x = g(w, x).$$

(In other words, action $w$ is the row $G_{w,-}$ in the matrix representation.)

- The expected gain of $w$ under prior $\pi$ is then given by the dot product

$$w \cdot \pi,$$

while $g$ can be seen as a set of such action vectors.

(We'll revisit this geometric representation ahead in this course.)

# Basic definitions

- $V_g$ is a measure of the vulnerability of threat to the secret.
- A complementary way measures the adversary's uncertainty about the secret.

  For that we can use **loss function**

$$\ell: \mathcal{W} \times \mathcal{X} \to \mathbb{R}_{\geq 0} \ ,$$

  assuming that the adversary wants to minimize her loss.

## Definition (3.4)

The $\ell$-**uncertainty** of $\pi$ is defined as

$$U_\ell(\pi) \ := \ \min_{w \in \mathcal{W}} \sum_{x \in \mathcal{X}} \pi_x \ell(w, x) \quad .$$

If $\mathcal{W}$ is infinite then the min should be replaced by inf.

- Although gain functions are allowed to take negative values, loss functions need to be non-negative. (Otherwise $U_\ell$ itself becomes negative).

# Basic definitions

- Note that uncertainties can be converted to vulnerabilities (and vice versa).

  By complementing

$$g = B - \ell \ ,$$

  where $B$ is some upper bound of $U_\ell$, we get

$$V_g = B - U_\ell \ .$$

# Basic definitions: Graphing $g$-vulnerability

- It's helpful to graph the vulnerability $V_g$ as a function of the prior $\pi$.

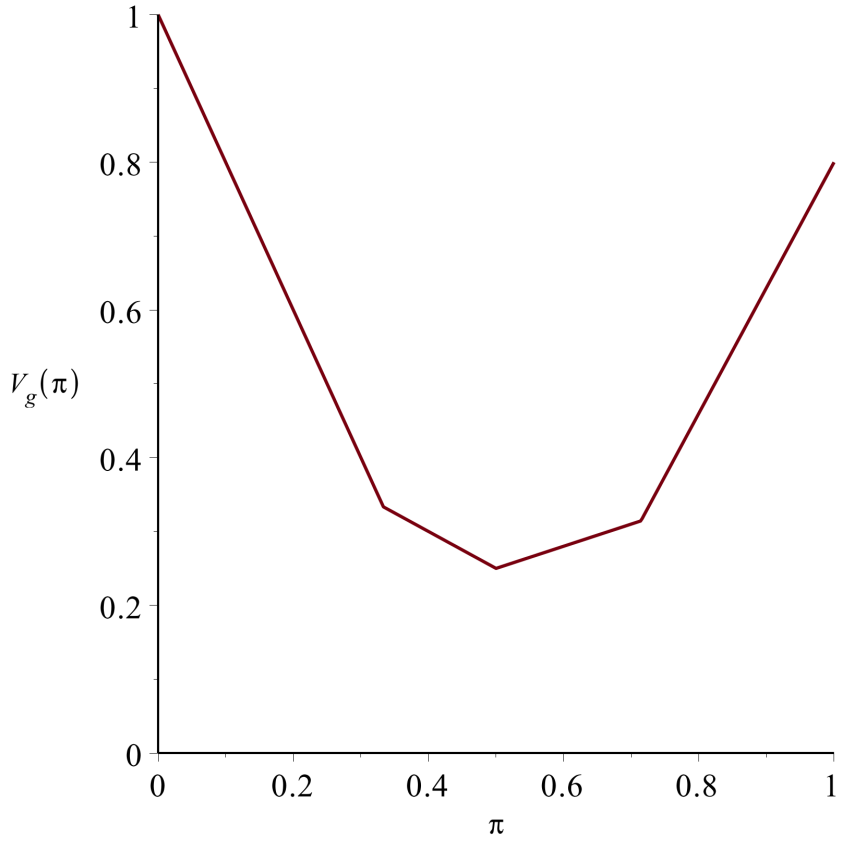
Example 2 Consider again the scenario having of a previous example, having:

- secret set $\mathcal{X} = \{x_1, x_2\}$;
- prior $\pi = (0.3, 0.7)$;
- action set $\mathcal{W} = \{w_1, w_2, w_3, w_4, w_5\}$; and
- gain function $g: \mathcal{W} \times \mathcal{X} \to \mathbb{R}$ having values as in

| G | $x_1$ | $x_2$ |
|---|---|---|
| $w_1$ | $-1.0$ | 1.0 |
| $w_2$ | 0.0 | 0.5 |
| $w_3$ | 0.4 | 0.1 |
| $w_4$ | 0.8 | $-0.9$ |
| $w_5$ | 0.1 | 0.2 |

.

We can graph the $V_g(\pi)$ as below:



# Basic definitions: Graphing $g$-vulnerability

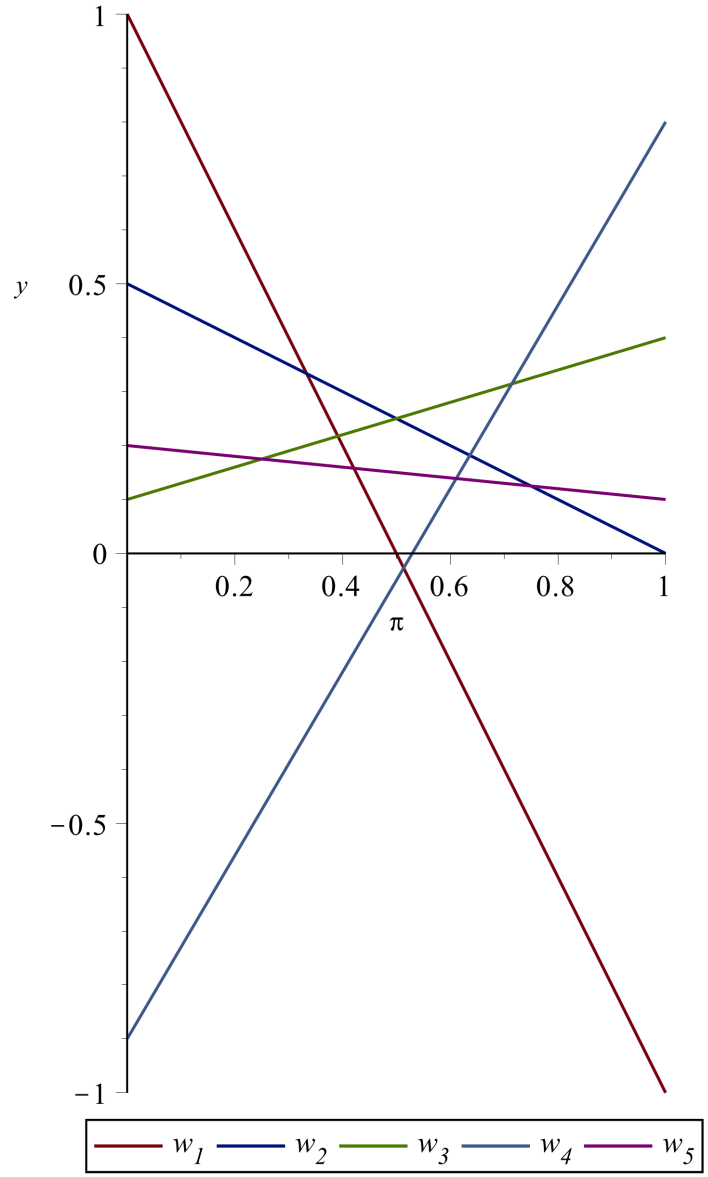
- In the example above, the graph of $V_g$ is:

  1. convex ($\smile$); and
  2. continuous.

  Those properties hold of $V_g$ for every gain function $g$ used.

- To understand why, consider the graphs of the expected gains provided by each of the five actions $w_1$, $w_2$, $w_3$, $w_4$, and $w_5$.

  This graph shows some important things about $g$-vulnerability.

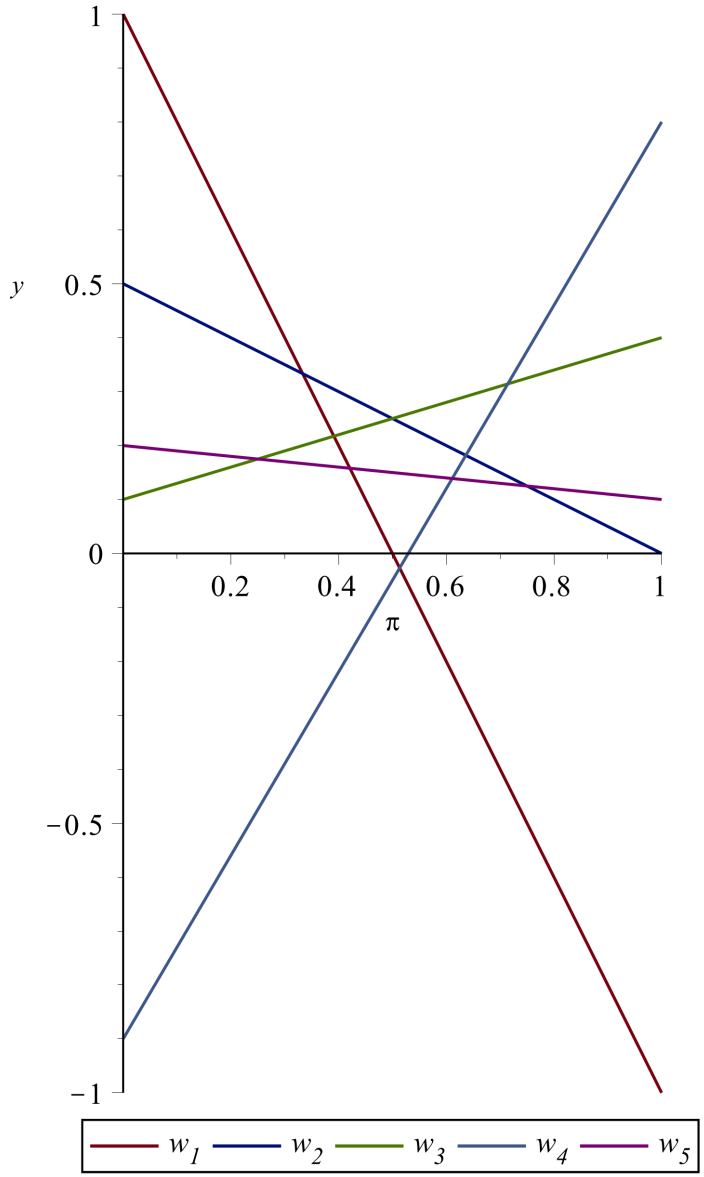


# Basic definitions: Graphing $g$-vulnerability

- The line of action $w_4$ connects the points $(0, -0.9)$ and $(1, 0.8)$.
- The graph of $V_g(\pi)$ lies along the maximum of all of the lines in the graph.

  $V_g$'s four line segments correspond to the four sets of distributions $\pi$ for which each action is optimal.
- The three interior vertices are the places where the adversary changes her mind about the best action.

  E.g., at the leftmost vertex, there's a tie between $w_1$ and $w_2$ (with $w_1$ being optimal to the left, and $w_2$ to the right).
- Action $w_5$ is sometimes better than other actions, but never better than all of them; an optimal adversary will never choose it.



# A catalog of gain functions

# A catalog of gain functions

- The gain function of the previous example was arbitrary.
- We'll now see a number of gain functions chosen more deliberately.

  They correspond to more realistic scenarios, and illustrate how $g$-vulnerability can model the threat to a secret in operationally significant ways.

# A catalog of gain functions: The identity gain function

- A reasonable gain function corresponds to a scenario where the adversary
  1. only by guessing the value of the secret exactly; and
  2. is allowed only one try.
- In this case:
  1. actions can simply be the values that can be guessed ($\mathcal{W}{=}\mathcal{X}$), and
  2. the gain function specifies the worth of correct and wrong guesses.

## Definition (3.5)

The **identity gain function** $g_{\mathsf{id}}\colon \mathcal{X}\times\mathcal{X} \to \{0,1\}$ is given by

$$g_{\mathsf{id}}(w,x) \;:=\; \begin{cases} 1, & \text{if } w = x, \\ 0, & \text{if } w \neq x \end{cases} .$$

# A catalog of gain functions: The identity gain function

- Note that $g_{\text{id}}$ takes $\mathcal{W}=\mathcal{X}$, so the adversary is required to make a guess.
- The matrix representation of $g_{\text{id}}$ is the identity matrix indexed by $\mathcal{X}\times\mathcal{X}$:

| $g_{\text{id}}$ | $x_1$ | $x_2$ | $x_3$ | $\ldots$ | $x_n$ |
|---|---|---|---|---|---|
| $x_1$ | 1 | 0 | 0 | $\ldots$ | 0 |
| $x_2$ | 0 | 1 | 0 | $\ldots$ | 0 |
| $x_3$ | 0 | 0 | 1 | $\ldots$ | 0 |
| $\ldots$ | 0 | 0 | 0 | $\ddots$ | 0 |
| $x_n$ | 0 | 0 | 0 | $\ldots$ | 1 |

.

# A catalog of gain functions: The identity gain function

- We can show that $g$-vulnerability is a generalization of Bayes vulnerability.

**Theorem (3.6)**

*Vulnerability under $g_{\text{id}}$ coincides with Bayes vulnerability: we have*

$$V_{g_{\text{id}}}(\pi) \quad = \quad V_1(\pi) \quad .$$

- **Proof.** Since

$$\sum_x \pi_x g_{\text{id}}(w, x) \;=\; \pi_w,$$

we have

$$V_{g_{\text{id}}}(\pi) \;=\; \max_w \pi_w \;=\; V_1(\pi).$$

□

# A catalog of gain functions: The identity gain function

- A natural extension of $g_{\text{id}}$ is to consider an adversary who:
  1. still benefits only by guessing the secret in one try, but
  2. different secrets lead to different gains.
- As we'll see later in this course, and interesting case is when the gain is defined as the **reciprocal** of the probability that some distribution $\sigma$ (i.e. not necessarily the prior) assigns to that secret.

## Definition (3.7)

For $\sigma : \mathbb{D}\mathcal{X}$, the **distribution-reciprocal gain function** $g_{\sigma^{-1}}$ is given by

$$\mathcal{W} = \lceil \sigma \rceil$$

and

$$g_{\sigma^{-1}}(w, x) \;:=\; \begin{cases} 1/\sigma_x & \text{if } w = x \\ 0 & \text{otherwise} \end{cases} .$$

# A catalog of gain functions: Gain functions induced from distance functions

- Sometimes the set $\mathcal{X}$ might come equipped with a **distance function**

$$d\colon (\mathcal{X}\times\mathcal{X}) \to \mathbb{R}_{\geq 0} \ ,$$

  indicating a natural notion of "distance" among secrets.

  (For instance, when secrets are geographic locations, or real numbers.)

- In such cases we can establish the gain function as a function of the distance.

  Given an additional function $h\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ of our choice, we can define a gain function $g_{d,h}$ **induced** by $d$ and $h$:

$$g_{d,h}(w,x) \ := \qquad h(d(w,x)) \quad .$$

  (Note that here we are taking $\mathcal{W} = \mathcal{X}$.)

- Given a distance $r$, we can make, for instance:

  (a) $h(r) = \max\limits_{x,x' \in \mathcal{X}} d(x,x') - r$

  (b) $h(r) = 1/r$

  (c) $h(r) = e^{-r}$

  (d) ...

# A catalog of gain functions: Gain functions induced from distance functions

- Gain functions induced from metrics are very expressive.
- Example 3 The identity gain function $g_{\text{id}}$ (i.e. Bayes vulnerability) is induced by the **discrete metric**,

$$d(w,x) \;:=\; \begin{cases} 0 & \text{if } w = x \\ 1 & \text{otherwise} \end{cases} ,$$

with

$$g_{\text{id}}(w,x) = 1 - d(w,x) \ ,$$

because it assigns distance 1 to any two distinct values. ◁

# A catalog of gain functions: Gain functions induced from distance functions

- Sometimes it's useful to allow the set $\mathcal{W}$ of guesses to be a superset of $\mathcal{X}$.

- Example 4 Suppose that the space of secrets is the set of corner points of a unit square: thus $\mathcal{X} = \{(0,0),(0,1),(1,0),(1,1)\}$.

  Suppose further that we use the gain function

$$g_{d,h}(w,x) = \sqrt{2} - d(w,x),$$

  where

$$d(w,x) = \|w - x\|_2$$

  is the Euclidean distance, and

$$h(r) = \sqrt{2} - r.$$

# A catalog of gain functions: Gain functions induced from distance functions

- Example 4 (Continued)

For a uniform prior, it is easy to see that any of the four corner points are equally good guesses, giving

$$V_{g_{d,h}}(\vartheta) \quad = \quad \frac{1}{4}\left(\sqrt{2} + 2(\sqrt{2}-1) + 0\right) \quad \approx \quad 0.5606 \quad .$$

But the adversary could actually do better by guessing

$$(1/2, 1/2),$$

a value that is not in $\mathcal{X}$.

Indeed, since that guess has distance $\sqrt{1/2}$ from each of the four corner points, we get a more favorable result for the adversary:

$$V_{g_{d,h}}(\vartheta) = \sqrt{1/2} \approx 0.7071.$$

# A catalog of gain functions: Gain functions induced from distance functions

- Also non-symmetric distances can be sometimes appropriate.

- Example 5 Suppose that the secret is the time (rounded to the nearest minute) that the last RER-B train will depart from Lozère back to Paris.

  The adversary (i.e. the weary traveler) wants to guess this time as accurately as possible, but note that:

  - guessing 23h44 when the actual time is 23h47

    is completely different from

  - guessing 23h47 when the actual time is 23h44!

  If we normalize so that a wait of an hour or more is considered intolerable, then we would want the non-symmetric distance function:

$$d(w,x) \;:=\; \begin{cases} {}^{(x-w)}/_{60} & \text{if } x-60 < w \leq x \\ 1 & \text{otherwise} \end{cases},$$

# A catalog of gain functions: Binary gain functions

- We now focus on binary gain functions, which return either 0 or 1:
  1. each action corresponds to the subset of $\mathcal{X}$ for which that action gives gain 1;
  2. we can assume wlog. that no two actions correspond to the same subset of $\mathcal{X}$, since such actions might as well be merged into one.

  Hence we can use the subsets themselves as the actions.

## Definition (3.8)

Given $\mathcal{W} \subseteq 2^{\mathcal{X}}$ with $\mathcal{W}$ nonempty, the **binary gain function** $g_{\mathcal{W}}$ is given by

$$g_{\mathcal{W}}(W, x) \;:=\; \begin{cases} 1, & \text{if } x \in W \\ 0, & \text{otherwise} \end{cases} .$$

- We'll now see some interesting gain functions with different choices of $\mathcal{W}$.

# A catalog of gain functions: Binary gain functions

- Example 6 **Two-block gain functions.** If

$$\mathcal{W} = \{W, \mathcal{X} - W\},$$

then we can see $W$ as a property that the secret $X$ might or might not satisfy.

Then $g_{\mathcal{W}}$ is the gain function corresponding to an adversary that merely wants to decide whether or not $X$ satisfies that property. ◁

# A catalog of gain functions: Binary gain functions

- Example 7 **Partition gain functions.** More generally, $\mathcal{W}$ could be any partition of $\mathcal{X}$ into one or more disjoint blocks.

  Then the adversary wants to determine which block the secret belongs to.

  That's equivalent to saying that

  $$\mathcal{W} = \mathcal{X}/\sim,$$

  by which we mean the set of blocks induced by the quotient with respect to an equivalence relation $(\sim)$ on $\mathcal{X}$.

  There are two extremes:

  - If $(\sim)$ is the identity relation, then the elements of $\mathcal{W}$ are all singletons, which means that $g_\sim$ is essentially the same as $g_{\mathsf{id}}$.
  - If $(\sim)$ is the universal relation, then $\mathcal{W}$ consists of a single block: thus $\mathcal{W} = \{\mathcal{X}\}$.

◁

# A catalog of gain functions: Binary gain functions

- Example 8 **The $k$-tries gain function.** Consider the adversary is allowed to make $k$ guesses, rather than just 1.

  If we define

$$\mathcal{W}_k \ := \quad \{W \mid W \subseteq \mathcal{X} \wedge |W| \leq k\} \quad ,$$

  then $V_{g_{\mathcal{W}_k}}(\pi)$ has exactly the following operational interpretation:

  *"The probability that the adversary could guess the value of $X$ correctly within $k$ tries."*

  Notice that $g_{\mathcal{W}_k}$ is not a partition gain function if $k>1$, since in that case its blocks overlap. ◁

# A catalog of gain functions: Binary gain functions

- Example 9 **General binary gain functions** In a general binary gain function, the actions $\mathcal{W}$ can be any nonempty subset of $2^{\mathcal{X}}$.

  Each element of $\mathcal{W}$ can be understood as a property that $X$ might satisfy.

  Then $g_{\mathcal{W}}$ is the gain function of an adversary who wants to guess any of those properties that $X$ satisfies. $\triangleleft$

# A catalog of gain functions: Gain function for a password database

- We now consider a more concrete operational scenario.

- Example 10 **Gain function for a password database.** Suppose that the secret is an array $X$ containing 10-bit passwords for a set $U$ of 1000 users, so that $X[u]$ is user $u$'s password.

  What gain functions might be suitable here?

  One possibility is to use the identity gain function $g_{\mathsf{id}}$, which as we have seen corresponds to Bayes vulnerability.

  But that specifies that the adversary gains only by correctly guessing the entire array $X$, which is not appropriate if we view $X$ as consisting of 1000 secrets, each of which is individually valuable.

# A catalog of gain functions: Gain function for a password database

- Example 10 (Continued)

We can instead define a gain function $g_{\text{pw}}$ that describes an adversary who wants to guess some user's password, with no preference as to whose it is.

That suggests

$$\mathcal{W} \ := \ \{(u,x) \mid u \in U \ \wedge \ 0 \leq x < 1024\}$$

and

$$g_{\text{pw}}((u,x),X) \ := \ \begin{cases} 1 & \text{if } X[u] = x \\ 0 & \text{otherwise} \end{cases} .$$

# A catalog of gain functions: Gain function for a password database

- Example 10 (Continued)

To see the effect of $g_{\text{pw}}$, suppose that the prior distribution on $X$ is uniform, which means that the 1000 passwords are independent and uniformly distributed over those 10 bits. Hence:

- The expected gain of every action $(u, x)$ in $\mathcal{W}$ is $2^{-10}$, which means that

$$V_{g_{\text{pw}}}(\pi) = 2^{-10}.$$

- In contrast,

$$V_{g_{\text{id}}}(\pi) = 2^{-10,000}.$$

That matches our intuition that:

- under $g_{\text{pw}}$ the adversary just needs to guess any single 10-bit password, while
- under $g_{\text{id}}$ she needs to guess 1000 such passwords.

# A catalog of gain functions: Gain function for a password database

- Example 10 (Continued)

It's interesting to consider a variant of $g_{\text{pw}}$ that:

1. allows the adversary to guess just a password $x$, without specifying whose it is, and
2. gives a gain of 1 if $x$ is correct for any of the users.

That suggests

$$\mathcal{W}' := \{x \mid 0 \leq x < 1024\}$$

and

$$g_{\text{pw}'}(x, X) := \begin{cases} 1 & \text{if } X[u] = x \text{ for some } u \in U \\ 0 & \text{otherwise} \end{cases}.$$

# A catalog of gain functions: Gain function for a password database

- Example 10 (Continued)

  Again using the uniform prior $\pi$, we now find that

$$V_{g_{\mathsf{pw}'}}(\pi) \quad = \quad 1 - \left({}^{1023}/_{1024}\right)^{1000} \quad \approx \quad 0.6236 \quad ,$$

  since that is the probability that any given $x$ occurs somewhere in $X$.

  But $g_{\mathsf{pw}'}$ does not seem a very reasonable gain function, since really a password is valuable only with respect to a particular user.

# A catalog of gain functions: Gain function for a password database

- Example 10 (Continued)

As yet another possibility, we can consider a gain function that specifies that some passwords are more valuable than others.

In particular, suppose that one of the users is Bill Gates, whose password is vastly more valuable than all the other passwords.

Then it would be appropriate to replace $g_{\text{pw}}$ with a gain function like

$$g_{\text{Gates}}((u,x),X) \;:=\; \begin{cases} 1 & \text{if } u = \text{Bill Gates and } x = X[u] \\ 0.01 & \text{if } u \neq \text{Bill Gates and } x = X[u] \\ 0 & \text{otherwise} \quad . \end{cases}$$

# A catalog of gain functions: Gain function for a password database

- Example 10 (Continued)

  With respect to a uniform prior $\pi$, we find that

  $$V_{g_{\text{Gates}}}(\pi) = 2^{-10},$$

  since that is the expected gain for any guess (Bill Gates, $x$) where $0 \leq x < 1024$.

  (Making a guess about any $u$ other than Bill Gates has an expected gain only one-hundredth as large.)

# A catalog of gain functions: Gain function for a password database

- Example 10 (Continued)

Note that $V_{g_{\text{Gates}}}$ isn't invariant w.r.t. permutations of the probabilities in $\pi$.

Consider a distribution $\pi^{u,x}$ where some user $u$'s password is known to be $x$, and all others are uniform and independent:

$$\pi^{u,x}_X := \begin{cases} 2^{-9990} & \text{if } X[u] = x \\ 0 & \text{otherwise} \end{cases} .$$

Different values of $u$ and $x$ simply permute the probabilities assigned by $\pi^{u,x}$.

Note that this has no effect on:

- The Shannon entropy: $\forall u, x,\ H(\pi^{u,x}) = 9990$ bits.
- The Bayes vulnerability, $\forall u, x,\ V(\pi^{u,x}) = 2^{-9990}$.

But it affects $V_{g_{\text{Gates}}}$:

- $V_{g_{\text{Gates}}}(\pi^{\text{Bill Gates},x}) = 1$,
- for $u \neq$ Bill Gates, $V_{g_{\text{Gates}}}(\pi^{u,x}) = 0.01$.

# A catalog of gain functions: Gain function for a password database

- Example 10 (Continued)

$V_{g_{\text{pw}}}$ is not invariant with respect to permutations of probabilities either.

Consider a distribution $\pi$ in which 10 of the bits in $X$ are known.

- If the 10 known bits all come from the same password, then

$$V_{g_{\text{pw}}}(\pi) = 1.$$

- But if the 10 known bits come from 10 different passwords, then

$$V_{g_{\text{pw}}}(\pi) = 2^{-9},$$

since the adversary's best action is to try to guess the unknown 9 bits from any one of those 10 passwords.

This means that $g_{\text{pw}}$ captures the structure of the password database, where certain sets of bits are worth more than others. ◁

# A catalog of gain functions: Gain function for a password database

- Example 11 Consider again the previous example on a gain function for a password database, in which we provided the numeric values of information measures on various prior distributions.

  Expand these computations to verify these results, using the proper definition of a $g$-vulnerability: $V_g(\pi) = \max_{w \in \mathcal{W}} \sum_{x \in \mathcal{X}} \pi_x \, g(w, x)$.

  (a) For the uniform prior $\pi$:

  (i) $V_{g_{\text{pw}}}(\pi)$ (ii) $V_{g_{\text{id}}}(\pi)$ (iii) $V_{g_{\text{pw}'}}(\pi)$ (iv) $V_{g_{\text{Gates}}}(\pi)$

  (b) For the prior $\pi_X^{u,x}$:

  (i) $V_{g_{\text{Gates}}}(\pi^{\text{Bill Gates},x})$ (ii) $V_{g_{\text{Gates}}}(\pi^{u,x})$, when $u \neq$ Bill Gates.

  (c) $V_{g_{\text{pw}}}(\pi)$ for prior $\pi$ in which 10 bits are known, all from the same password.

  (d) $V_{g_{\text{pw}}}(\pi)$ for prior $\pi$ in which 10 bits are known, each from a different password.

  **Solution.** For the student! ◁

# A catalog of gain functions: A gain function that penalizes wrong guesses

- Example 12 Now imagine a scenario where the adversary tries to input an access code $X$ into a keypad outside a locked door:
  1. inputting the correct code unlocks the door, but...
  2. inputting the wrong code triggers a penalty (say, opening a trap door to a pit of tigers).

  This scenario can be modeled with a gain function $g_{\text{tiger}}$ using

$$\mathcal{W} = \mathcal{X} \cup \{\bot\},$$

  where the special action $\bot$ is used to opt not to input a guess, and

$$g_{\text{tiger}}(w, x) \;:=\; \begin{cases} 1 & \text{if } w = x \\ 0 & \text{if } w = \bot \\ -1 & \text{otherwise} \end{cases} .$$

◁

# A catalog of gain functions: A gain function for a medical diagnosis scenario

- Example 13 Related to $g_{\text{tiger}}$ is a gain function for a scenario in which the adversary is trying to diagnose a certain disease.

  Here the set of possible values of the secret is

  $$\mathcal{X} = \{\textit{disease}, \textit{no disease}\} \ .$$

  Of course she would like to guess correctly whether or not the disease is present, but she is more interested in what action to take.

  Thus she might have

  $$\mathcal{W} = \{\textit{treat}, \ \textit{don't treat}\} \ .$$

  In this example, we are inclined to view the "adversary" as benevolent.

  Hence here we want to foster, rather than not prevent, information flow!

# A catalog of gain functions: A gain function for a medical diagnosis scenario

- Example 13 (Continued)

  Or she might have a set of possible treatments, some aggressive and some conservative.

  What is interesting here is that, unlike in the case of $g_{\text{tiger}}$, different errors would bring different penalties.

  Thus we might want a gain function something like

| $g_{\text{diagnosis}}$ | *disease* | *no disease* |
|---|---|---|
| *treat* | 5 | $-2$ |
| *don't treat* | $-5$ | 2 |

◁

# A catalog of gain functions: A loss function that gives guessing entropy

- To get a sense of the expressiveness of the $g$-vulnerability framework, we now consider how we can express guessing entropy.
- The **guessing entropy** of a secret $X$ with distribution $\pi$ is defined as the expected number of brute-force tries needed to guess the secret.

  The adversary's best strategy is to try possible values in non-increasing order of probability.

  If $x_n$ is an indexing of $\mathcal{X}$ in such an order, and $|\mathcal{X}|=N$, then the guessing entropy is defined by

$$G(\pi) \;:=\; \sum_{n=1}^{N} n\pi_{x_n} \quad .$$

# A catalog of gain functions: A loss function that gives guessing entropy

- The operational interpretation of guessing entropy is the following.
  1. the adversary must make brute-force guesses of the form

     *"Is the secret equal to $x_n$?"*;

  2. the loss is the number of guesses needed to find the correct value;
  3. the actual action corresponding to a permutation $f$ is the "strategy" of guessing the values of $\mathcal{X}$ in the particular *order* specified by $f$.

# A catalog of gain functions: A loss function that gives guessing entropy

- Since guessing entropy is an uncertainty measure, we express it as an $\ell$-uncertainty, for some loss function $\ell$.

  We let the set of actions $\mathcal{W}$ contain all *permutations* $f$ of $\mathcal{X}$, and define

$$\ell_G(f, x) \;:=\; \qquad n \quad ,$$

  where $n$ is the unique index of $x$ within permutation $f$, (in range 1 to $N$).

  The expected loss from permutation $f$ is

$$\sum_{n=1}^{N} \pi_{f_n} \cdot n \quad ,$$

  which is minimized if $f$ orders $\mathcal{X}$ in non-increasing probability order.

  Hence we have

$$U_{\ell_G}(\pi) \quad = \quad G(\pi) \quad .$$

# A catalog of gain functions: A loss function that gives Shannon entropy

- We can also express Shannon entropy in the $g$-vulnerability framework if:
  1. we use a loss- (rather than gain-) function; and
  2. we use an infinite set $\mathcal{W}$ of actions.

  That is all explained in the book in Section 3.2.8!

# Classes of gain functions

# Classes of gain functions

- Now we turn our attention from particular gain functions to four classes of gain functions with properties that we will find useful subsequently.

| Class | Name | Description |
|---|---|---|
| $\mathbb{G}\mathcal{X}$ | **Finite-valued, non-negative vulnerabilities** | All gain functions $g$ such that $V_g$ is in $\mathbb{D}\mathcal{X} \to \mathbb{R}_{\geq 0}$. |
| $\mathbb{G}^{\text{fin}}\mathcal{X}$ | **Finitely many actions** | All gain functions $g\colon \mathbb{G}\mathcal{X}$ having a finite set of actions $\mathcal{W}$. |
| $\mathbb{G}^{+}\mathcal{X}$ | **Non-negative gain functions** | All gain functions $g\colon \mathbb{G}\mathcal{X}$ that are themselves $\geq 0$. |
| $\mathbb{G}^{\updownarrow}\mathcal{X}$ | **One-bounded gain functions** | All gain functions $g\colon \mathbb{G}\mathcal{X}$ that are themselves $\leq 1$. This class coincides all $g$ such that $V_g$ is in $[0, 1]$. |

# Mathematical properties

# Mathematical properties

- We'll now explore some general mathematical properties of $g$-vulnerability.

## Definition

1. A vector $\sum_n a_n x^n$, where $x^1, \ldots, x^N$ are vectors and $a_1, \ldots, a_N$ are non-negative reals adding up to 1, is called a **convex combination** (and the $a_n$'s are called **convex coefficients**).

2. A **set is convex** iff it contains all convex combinations of its elements.

3. A **function** $f$, defined on a convex set, is called **convex** iff

$$f\left(\sum_n a_n x^n\right) \leq \sum_n a_n f(x^n),$$

and **linear** iff

$$f\left(\sum_n a_n x^n\right) = \sum_n a_n f(x^n).$$

- Note that it is easy to see that $\mathbb{D}\mathcal{X}$ is a convex set.

# Mathematical properties

- We have the following theorem.

## Theorem (3.13)

*For any $g\colon \mathbb{G}\mathcal{X}$, the induced $V_g(\pi)$ is a convex function of $\pi$.*

**Proof.** Note that $V_g$ can be expressed as the sup of the convex (and in fact linear) functions $\pi \mapsto \sum_x \pi_x g(w,x)$; and it is well known that the sup of convex functions is convex.

To show the theorem directly in the case $g\colon \mathbb{G}^{\text{fin}}\mathcal{X}$, we have

$$
\begin{array}{lll}
 & V_g\left(\sum_i a_i \pi^i\right) & \\
= & \max_w \sum_x (\sum_i a_i \pi^i)_x \; g(w,x) & \\
= & \max_w \sum_x \sum_i a_i \pi^i_x \; g(w,x) & \\
= & \max_w \sum_i a_i \sum_x \pi^i_x \; g(w,x) & \\
\leq & \sum_i a_i \max_w \sum_x \pi^i_x \; g(w,x) & \\
= & \sum_i a_i V_g(\pi^i) & .
\end{array}
$$

# Mathematical properties

- The intuition behind the convexity of $V_g$ is that
  - in

$$V_g\left(\sum_i a_i \pi^i\right)$$

    the adversary is forced to select a single action for the "composite" distribution $\sum_i a_i \pi^i$, while
  - in

$$\sum_i a_i V_g(\pi^i)$$

    the adversary is able to select a different, "custom" action for each $\pi^i$, potentially enabling a better expected gain.

# Mathematical properties: Gain-function algebra

- Given a gain function $g: \mathcal{W} \times \mathcal{X} \to \mathbb{R}$:
  - the particular values returned by $g$ are not very important;
  - what really matters is how the various gain values compare with one another.

  Hence, modifying the gain values in certain ways leads to predictable modifications of $g$-vulnerabilities.

- We consider two modifications:
  - **Scaling**, where we multiply all gain values by a non-negative constant $k$.

    That is, we define a **scaled** gain function:

$$g_{\times k}(w, x) \;:=\; k \times g(w, x) \quad .$$

  - **Shifting**, where we add a value $\kappa_x$, possibly depending on $x$, to all gain values.

    That is, we define a **shifted** gain function:

$$g_{+\kappa}(w, x) \;:=\; g(w, x) + \kappa_x \quad .$$

# Mathematical properties: Gain-function algebra

- The choice of the vector $\kappa$ gives us flexibility in the way the shifting is performed.

  For instance:

  (a) We could shift all gain values by the same amount $k$ by choosing $\kappa = k1$;

  (b) We could shift the gain values of a single secret by choosing $\kappa = k[x]$.

  Viewing $g$ as a set of gain vectors $w$, we see that scaling and shifting are simply the scalar multiplication $kw$ and vector addition $w + \kappa$ of all gain vectors respectively.

  Similarly, when $g$ is represented as a matrix G, scaling consists of multiplying G by $k$, while shifting consists of adding $\kappa$ to all rows of G.

# Mathematical properties: Gain-function algebra

- The next result shows how the vulnerabilities of a modified gain function can be derived from the vulnerabilities of the original gain function.

## Theorem (3.14)

*For any $g\colon \mathbb{G}\mathcal{X}$, prior $\pi$, $k \geq 0$ and $\kappa \in \mathbb{R}^{|\mathcal{X}|}$ the following hold:*

$$\begin{aligned} V_{g_{\times k}}(\pi) &= k V_g(\pi)\ , \\ V_{g_{+\kappa}}(\pi) &= V_g(\pi) + \kappa \cdot \pi \quad . \end{aligned}$$

**Proof.** Using the representation of $g$ as a set of action vectors, we reason as

$$\begin{aligned} & V_{g_{\times k}}(\pi) \\ = \quad & \sup_{w \in g_{\times k}} w \cdot \pi && \text{“Def. of } V_g \text{, for action vectors”} \\ = \quad & \sup_{w \in g} kw \cdot \pi && \text{“} g_{\times k} = \{kw \mid w \in g\} \text{”} \\ = \quad & \sup_{w \in g} k(w \cdot \pi) \\ = \quad & k \sup_{w \in g} w \cdot \pi && \text{“} \sup_i kx_i = k \sup_i x_i \text{ for } k \geq 0 \text{”} \\ = \quad & k V_g(\pi) \end{aligned}$$

# Mathematical properties: Gain-function algebra

- **Proof.** (Continued)

  And

$$
\begin{array}{lll}
 & V_{g_{+\kappa}}(\pi) & \\
= & \sup_{w \in g_{+\kappa}} w \cdot \pi & \text{“Def. of } V_g\text{”} \\
= & \sup_{w \in g} (w + \kappa) \cdot \pi & \text{“} g_{+\kappa} = \{w + \kappa \mid w \in g\}\text{”} \\
= & \sup_{w \in g} (w \cdot \pi + \kappa \cdot \pi) & \\
= & (\sup_{w \in g} w \cdot \pi) + \kappa \cdot \pi & \text{“} \sup_i (x_i + c) = (\sup_i x_i) + c\text{”} \\
= & V_g(\pi) + \kappa \cdot \pi \quad . &
\end{array}
$$

  □

- An issue that must be considered, however, is the effect of scaling and shifting on membership in $\mathbb{G}\mathcal{X}$.

  For any $g$ in $\mathbb{G}\mathcal{X}$, it is easy to see that $g_{\times k}$ is also in $\mathbb{G}\mathcal{X}$ for any $k \geq 0$.

  But $g_{+\kappa}$ might *not* be in $\mathbb{G}\mathcal{X}$, since negative entries of $\kappa$ could make $V_{g_{+\kappa}}$ negative on some priors.

# On “absolute” versus “relative” security

# On "absolute" versus "relative" security

- Gain functions and $g$-vulnerability enable us to measure security in a rich variety of ways.

  But it's often hard to know precisely which gain function is most appropriate!

- However, often we:
  - don't care about the precise value of $V_g(\pi)$ ("absolute" security), but...
  - only care about how the expected gains of the various actions in $\mathcal{W}$ compare with each other ("relative" security).

- A feature of the proof of Theorem 3.14 is that the modified gain functions $g_{\times k}$ and $g_{+\kappa}$ always have the same optimal actions as does the original gain function $g$.

  Thus the modifications affect the precise vulnerability values, but they do not affect what action the adversary should take.

# On "absolute" versus "relative" security

- Example 14 Consider the case of a neighborhood under threat of burglary:
  - The secret $X$ consists in:
    - the contents of the various houses (e.g. jewelry, cash, electronics),
    - the security measures protecting them (e.g. locks, safes, alarms, surveillance cameras), and
    - time-dependent information (who's at home, awake, or sleeping, in the various houses).
  - The prior distribution $\pi$ models the adversary's (partial) knowledge of $X$.
  - The gain function $g$ models the actions $\mathcal{W}$ that the adversary could take.

    (E.g. break the glass door at the back of a certain house at a certain time.)
  - The gain values $g(w, x)$ model actions' results.

    (E.g. successful theft of a diamond necklace worth \$10,000 –good for her– or detection and arrest by police — bad).

Clearly here the complete scenario ($X$, $\pi$, and $g$) could be very complex, and not knowable precisely.

# On "absolute" versus "relative" security

- Example 14 (Continued)

  But if we are a homeowner in the neighborhood, our own goal is considerably more modest.

  When we think about what security devices we might buy for our house, our goal is not to make our house absolutely secure (which is in fact impossible), but only to make our house less tempting to the burglar than our neighbors' houses are — so that the burglar's optimal action is to go to one of their houses rather than ours.

  The particular gain values that lead to that don't really matter to us! ◁

# Appendix: More on classes of gain functions

# Classes of gain functions

- Here we give more details about the classes of gain functions we consider in this course.

| Class | Name | Description |
|---|---|---|
| $\mathbb{G}\mathcal{X}$ | **Finite-valued, non-negative vulnerabilities** | All gain functions $g$ such that $V_g$ is in $\mathbb{D}\mathcal{X} \to \mathbb{R}_{\geq 0}$. |
| $\mathbb{G}^{\text{fin}}\mathcal{X}$ | **Finitely many actions** | All gain functions $g\colon \mathbb{G}\mathcal{X}$ having a finite set of actions $\mathcal{W}$. |
| $\mathbb{G}^{+}\mathcal{X}$ | **Non-negative gain functions** | All gain functions $g\colon \mathbb{G}\mathcal{X}$ that are themselves $\geq 0$. |
| $\mathbb{G}^{\updownarrow}\mathcal{X}$ | **One-bounded gain functions** | All gain functions $g\colon \mathbb{G}\mathcal{X}$ that are themselves $\leq 1$. This class coincides all $g$ such that $V_g$ is in $[0, 1]$. |

# Classes of gain functions: Finite-valued, non-negative vulnerabilities - the class $\mathbb{G}\mathcal{X}$

- The most general class that we consider is $\mathbb{G}\mathcal{X}$, which is the set of all gain functions $g$ such that the induced $V_g$ gives *finite* and *non-negative* values.

**Definition (3.10)**

The class $\mathbb{G}\mathcal{X}$ consists of all gain functions $g$ such that $V_g$ is in $\mathbb{D}\mathcal{X} \to \mathbb{R}_{\geq 0}$.

- This class $\mathbb{G}\mathcal{X}$ requires:
  - **Finiteness of** $g$: it must be bounded from above.
  - **Non-negativity of** $V_g$: but it is still possible that $g$ has negative values.
- All gain functions considered earlier in our catalog belong to $\mathbb{G}\mathcal{X}$.
- Important: In this course we restrict our attention exclusively to gain functions in $\mathbb{G}\mathcal{X}$ (although we often restrict further to subsets of it).

# Classes of gain functions: Finitely many actions $\mathbb{G}^{\text{fin}}\mathcal{X}$

- To simplify the presentation, in many parts of this course we restrict to the class

$$\mathbb{G}^{\text{fin}}\mathcal{X} \subset \mathbb{G}\mathcal{X}$$

of **gain functions having a finite set of actions** $\mathcal{W}$.

- Note that any such gain function $g \in \mathbb{G}^{\text{fin}}\mathcal{X}$ is automatically bounded, making $V_g$ finite-valued.

  But we still require $V_g$ to be non-negative.

- The use of $\mathbb{G}^{\text{fin}}\mathcal{X}$ is implied whenever a gain function is expressed as a matrix G, or whenever $V_g$ is expressed as a max instead of a sup.

# Classes of gain functions: Non-negative gain functions $\mathbb{G}^{+}\mathcal{X}$

- The next restriction that we consider here is to require the gain function $g$ itself to be non-negative.

**Definition (3.11)**

The class $\mathbb{G}^{+}\mathcal{X}$ consists of all **non-negative gain functions** $g$ in $\mathbb{G}\mathcal{X}$.

- This restriction ensures that $V_g$ is also non-negative, but we still need $g$ to be bounded in order for $V_g$ to be finite-valued.

# Classes of gain functions: One-bounded gain functions $\mathbb{G}^{\updownarrow}\mathcal{X}$

- Finally we consider gain functions $g$ that are bounded from above by 1.

**Definition (3.12)**

The class $\mathbb{G}^{\updownarrow}\mathcal{X}$ consists of all gain functions $g\colon \mathbb{G}\mathcal{X}$ such that $g \leq 1$.

- Note that:
  - $g$ need not be bounded from below (except that, as required by $\mathbb{G}\mathcal{X}$, the induced $V_g$ must be non-negative).
  - if $g$ is one-bounded then $V_g$ is also one-bounded.

    And the converse holds as well: if $g(w, x) > 1$ for some $w$ and $x$, then $V_g([x]) > 1$.

  Hence $\mathbb{G}^{\updownarrow}\mathcal{X}$ can equivalently be defined as the set of all gain functions $g$ such that the induced vulnerability $V_g$ is always between 0 and 1.