Methods of Artificial Intelligence: Lecture

8. Session: Vagueness and Uncertainty I

Kai-Uwe Kühnberger, Nohayr Muhammad

Winter Term 2022/2023

December 23rd, 2022

Overview

- Remaining from Last Time: Description Logics - Inferences
- Fuzzy Logic: Introduction / Motivation
- Fuzzy Logic: Fuzzy Set Theory
- Fuzzy Logic: t-norms and s-norms
- Probabilistic Logic

Description Logics: Inferences

The Tableaux Algorithm

Inference Algorithms

- Two types of algorithms
  - Structural subsumption algorithms (for weak DLs)
  - Tableau-based algorithms (general technique)
- Remarks:
  - Relation of DLs to 2-variable logic
  - Most DLs can be reduced to 2-variable logic
  - Problematic cases are role composition and number restrictions: these operations cannot be expressed by 2-variable logic in general (why?)

Inference Algorithms

- Structural subsumption algorithms try to test subsumption of concept descriptions
  - This works only if no disjunction is available
  - Compare Baader & Nutt: “Basic Description Logic”

- Tableau-based algorithms reduce subsumption to the unsatisfiability of concept descriptions:

  $C \sqsubseteq D$ iff $C \sqcap \neg D$ is unsatisfiable
  (by checking if a (finite) model exists)

- We explain Tableau-based algorithms using an example
  - Assume we want to know whether $(\exists R.A) \sqcap (\exists R.B)$ is subsumed by $\exists R.(A \sqcap B)$
  - We must check whether
    $$C = (\exists R.A) \sqcap (\exists R.B) \sqcap \neg(\exists R.(A \sqcap B))$$
    is unsatisfiable

Inference Algorithms

- Tableau-based algorithms: an example
  - Check for $(\exists R.A) \sqcap (\exists R.B) \sqsubseteq \exists R.(A \sqcap B)$
  - We must check whether

$$C = (\exists R.A) \sqcap (\exists R.B) \sqcap \neg(\exists R.(A \sqcap B))$$

is unsatisfiable

- Push all negations as far as possible into the description (negation normal form)
  - $C' = (\exists R.A) \sqcap (\exists R.B) \sqcap (\forall R.(\neg A \sqcup \neg B))$
  - Assume that there exists a $b^I \in (C')^I$
  - This corresponds to finding a model for an A-box: $b : C'$

Inference Algorithms

Try to construct a finite interpretation $\mathcal{I}$ such that $(C')^{\mathcal{I}} \neq \emptyset$

1.	$b^{\mathcal{I}} \in (C')^{\mathcal{I}}$	
2.	$b^{\mathcal{I}} \in ((\exists R.A) \sqcap (\exists R.B) \sqcap (\forall R.(\neg A \sqcup \neg B)))^{\mathcal{I}}$	(1., def)
3.	$b^{\mathcal{I}} \in (\exists R.A)^{\mathcal{I}}$	(2., $\sqcap$)
4.	$b^{\mathcal{I}} \in (\exists R.B)^{\mathcal{I}}$	(2., $\sqcap$)
5.	$b^{\mathcal{I}} \in (\forall R.(\neg A \sqcup \neg B))^{\mathcal{I}}$	(2., $\sqcap$)
6.	$<b^{\mathcal{I}},c^{\mathcal{I}}> \in R^{\mathcal{I}}$	(3., skolemization)
7.	$c^{\mathcal{I}} \in A^{\mathcal{I}}$	
8.	$<b^{\mathcal{I}},d^{\mathcal{I}}> \in R^{\mathcal{I}}$	(4., skolemization)
9.	$d^{\mathcal{I}} \in B^{\mathcal{I}}$	
10.	$c^{\mathcal{I}} \in (\neg A \sqcup \neg B)^{\mathcal{I}}$	(5., 6., $\forall$)
11.	$d^{\mathcal{I}} \in (\neg A \sqcup \neg B)^{\mathcal{I}}$	(5., 8., $\forall$)
12.	$c^{\mathcal{I}} \in (\neg B)^{\mathcal{I}}$	(10., $c^{\mathcal{I}} \in (\neg A)^{\mathcal{I}}$ clashes with 7.: $c^{\mathcal{I}} \in A)^{\mathcal{I}}$
13.	$d^{\mathcal{I}} \in (\neg A)^{\mathcal{I}}$	(11., $d^{\mathcal{I}} \in (\neg B)^{\mathcal{I}}$ clashes with 9.: $d^{\mathcal{I}} \in B)^{\mathcal{I}}$

$\Delta^{\mathcal{I}} = \{b^{\mathcal{I}},c^{\mathcal{I}},d^{\mathcal{I}}\}$, $R^{\mathcal{I}} = \{<b^{\mathcal{I}},c^{\mathcal{I}}>, <b^{\mathcal{I}},d^{\mathcal{I}}>\}$, $A^{\mathcal{I}} = \{c^{\mathcal{I}}\}$, $B^{\mathcal{I}} = \{d^{\mathcal{I}}\}$ is a finite model for $b : C'$

Inference Algorithms

- We found a model for $b : C'$
- $C' = (\exists R.A) \sqcap (\exists R.B) \sqcap (\forall R.(\neg A \sqcup \neg B))$ is satisfiable
- $C = (\exists R.A) \sqcap (\exists R.B) \sqcap \neg(\exists R.(A \sqcap B))$ is not unsatisfiable
- $(\exists R.A) \sqcap (\exists R.B) \sqsubseteq \exists R.(A \sqcap B)$ does not hold!

Fuzzy Logic: Introduction / Motivation

Vagueness and Its Modeling

Vagueness and Uncertainty

- In Classical Logics, statements are either true or false
- We cannot capture
  - Partial Truth: the ball is reddish (the ball is red to a certain degree)

  - Uncertainty: we probably pick a red ball (most balls are red)

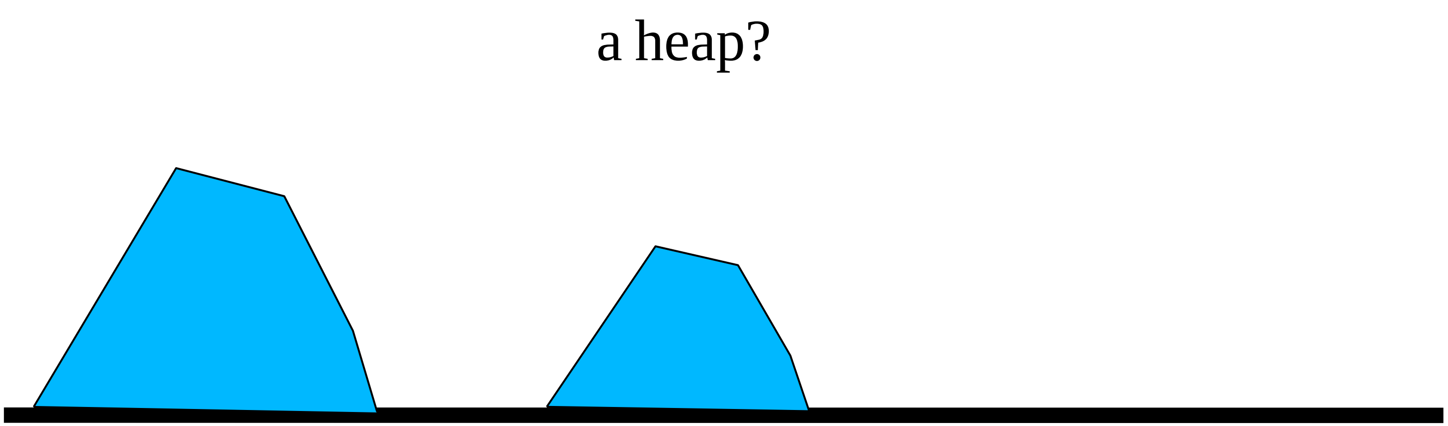
- Fuzzy Logics address partial truth (vagueness)
- Probabilistic Logics address uncertainty

The Heap Paradox

- 1,000,000 grains of sand is a heap of sand ***(Premise 1)***
- A heap of sand minus one grain is still a heap ***(Premise 2)***

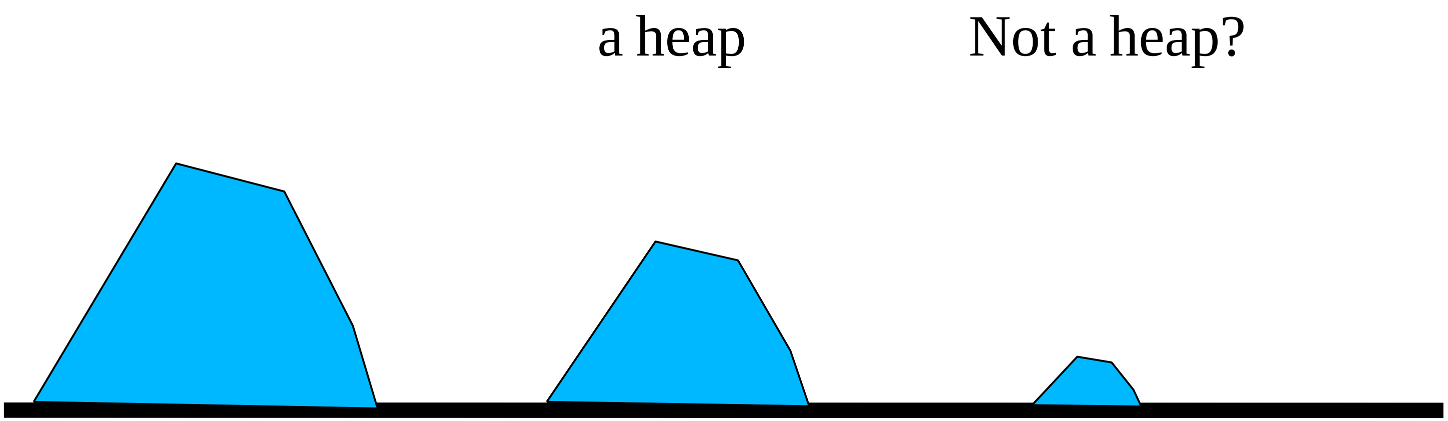
The Heap Paradox

- But then 999,999 is still a heap



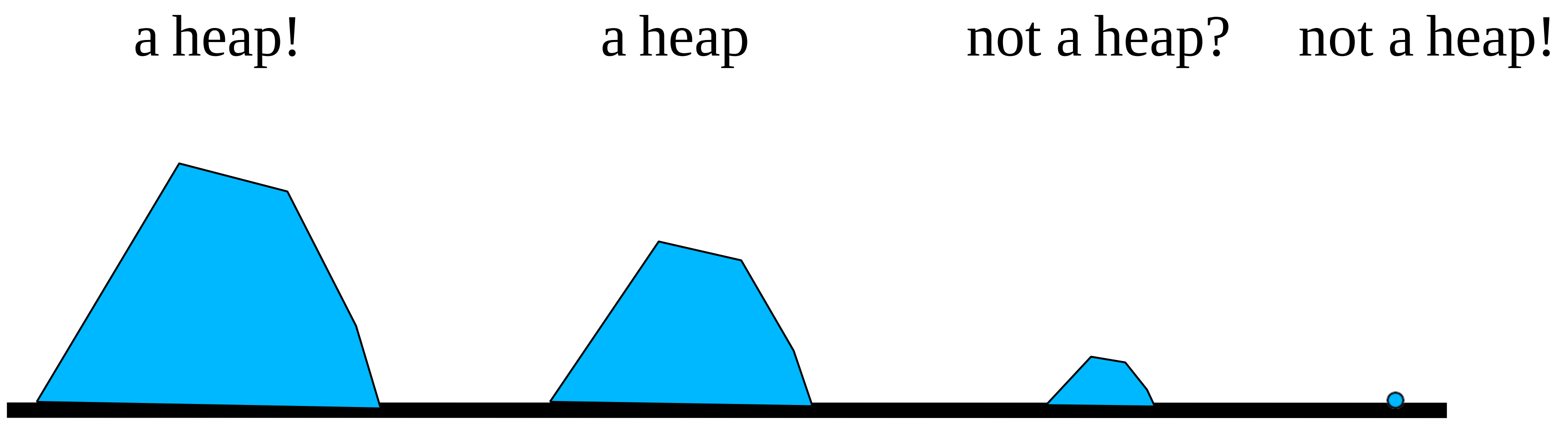
The Heap Paradox

- If we iterate the argument…



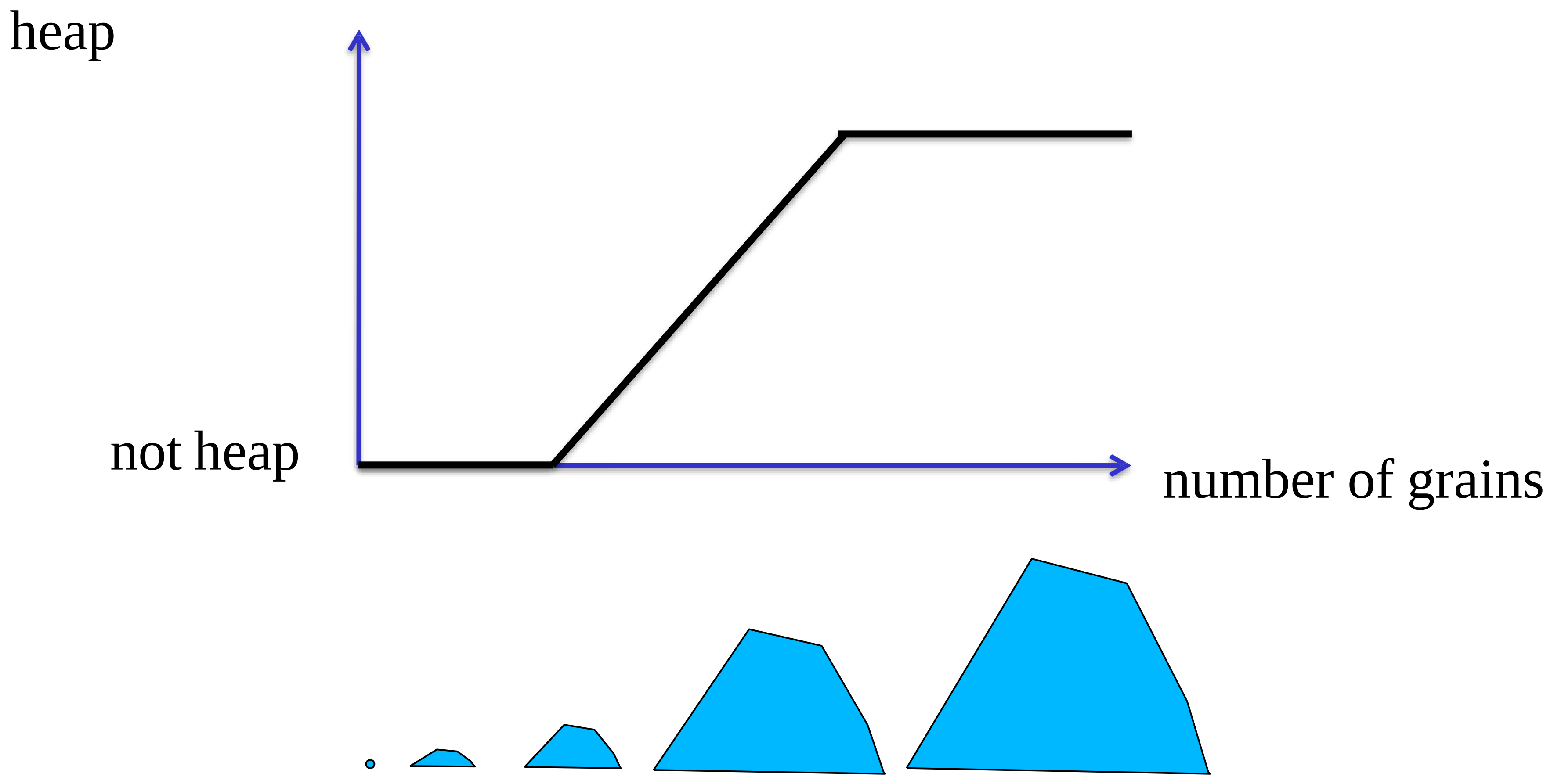
The Heap Paradox

- … we can conclude that 1 grain of heap is still a heap



Fuzzy Approach to the Heap Paradox

- Instead of classical Boolean truth values, quantify degree of truth by an arbitrary value between 0 and 1



Fuzzy Logic: Fuzzy Set Theory

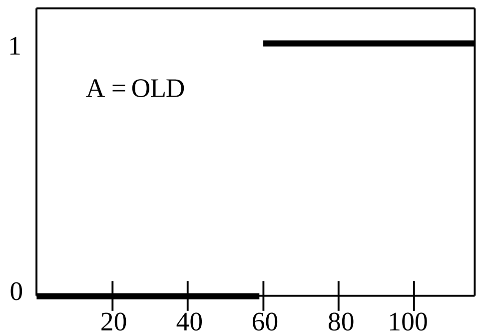
Modified Membership Relation

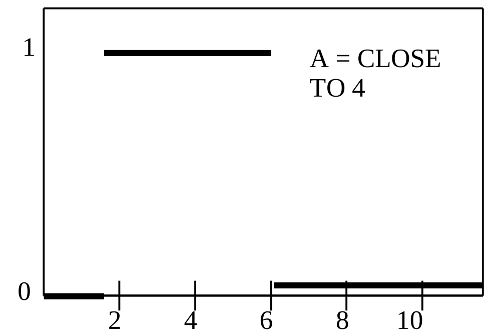
Membership Functions

- Classical sets can be represented by a characteristic function (also called indicator function)

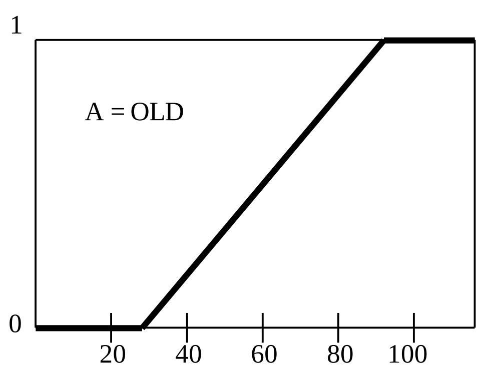
- A characteristic function yields 1 for all elements that belong to the set

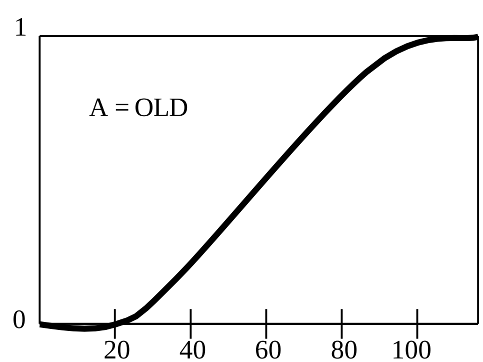
  $1_A(x) = 1 \quad iff \quad x \in A \quad (and\ 0\ otherwise)$

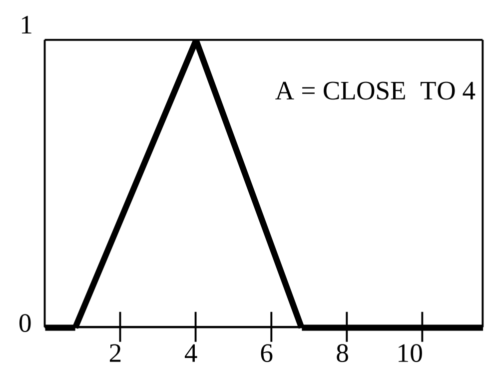




- Some fuzzy membership functions (*y*-value = degree of membership)







Fuzzy Set Theory

- A fuzzy set $A = (U, \mu_A)$ is defined by
  - some universe U and
  - a membership function $\mu_A : U \to [0, 1]$
- Intuitively, $\mu_A(x) = d$ means that x belongs to degree d to A
- For example, we could have $\mu_{red}(ball) = 0.6$

- Classical (crisp) sets are special membership functions $\mu_A: U \to \{0,1\}$
  - $\mu_A(x) = 1$ iff $x \in A$.
  - $\mu_A(x) = 0$ iff $x \notin A$.

(characteristic function)

Standard Fuzzy Set Theory

- We can generalize classical set relations
  - $A = B$ iff $\forall x.\ \mu_A(x) = \mu_B(x)$ (equality)
  - $A \subseteq B$ iff $\forall x.\ \mu_A(x) \leq \mu_B(x)$ (subsethood)

- One way to generalize set operations is as follows
  - $A \cap B : \mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$ (intersection)
  - $A \cup B : \mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$ (union)
  - $A^C : \mu_{A^C}(x) = 1 - \mu_A(x)$ (complement)

Example

- Suppose, we have fuzzy sets *red*, *round* s.t.
  - $\mu_{red}(\text{ball}) = 0.6$
  - $\mu_{round}(\text{ball}) = 1.0$

- Then, we have
  - For S = red $\cap$ round : $\mu_S(\text{ball}) = \min(0.6,1.0) = 0.6$
    - *Ball is both red and round to degree 0.6*
  - For S = red $\cup$ round : $\mu_S(\text{ball}) = \max(0.6,1.0) = 1.0$
    - *Ball is red or round (to degree 1.0; ~ classical truth)*
  - For S = red^C: $\mu_S(\text{ball}) = 1.0 - 0.6 = 0.4$
    - *Ball is not red to the degree 0.4*

Some Other Properties

- Using the previous definitions, we maintain many classical properties of sets
  - Neutral elements:
    - $\forall x.\ \min(1,\mu_A(x)) = \mu_A(x)$ $(U \cap A = A)$
    - $\forall x.\ \max(0,\mu_A(x)) = \mu_A(x)$ $(\emptyset \cup A = A)$
  - Commutativity:
    - $\forall x.\ \min(\mu_A(x), \mu_B(x)) = \min(\mu_B(x), \mu_A(x))$ $(A \cap B = B \cap A)$
    - $\forall x.\ \max(\mu_A(x), \mu_B(x)) = \max(\mu_B(x), \mu_A(x))$ $(A \cup B = B \cup A)$
  - Similarly, other properties remain true like
    - Associativity, distributivity, deMorgan's Laws, …

Fuzzy vs. Classical Sets

- The intersection of non-crisp fuzzy sets and their complement can be ‘non-empty’
  - For classical sets, we have

$$A \cap A^c = \varnothing$$

  - For fuzzy sets, we have

$$\mu_{A \cap A^c}(x) = \min(\mu_A(x), 1 - \mu_A(x))$$

  - Example: let S = red $\cap$ red^c
    Then $\mu_S(x) = \min(0.6, 1.0 - 0.6) = 0.4$

- Of course, this makes only sense for ‘vague predicates’

Fuzzy vs. Classical Sets

- The union of fuzzy sets and their complement can be different from the universe U
  - For classical sets, we have

$$A \cup A^c = U$$

  - For fuzzy sets, we have

$$\mu_{A \cup A^c}(x) = \max(\mu_A(x), 1 - \mu_A(x))$$

  - Example: let S = red $\cup$ red^C
    Then $\mu_S(x) = \max(0.6, 1.0 - 0.6) = 0.6$

- Of course, this makes only sense for 'vague predicates'

Fuzzy Logic

t-norms and s-norms

Fuzzy Logic

- We can define a propositional Fuzzy Logic as follows:
  - Interpretations I assign a membership value from [0,1] to all atoms
  - we interpret conjunctions using *min*:
    - $\mathrm{I}(F \wedge G) = min(\mathrm{I}(F),\mathrm{I}(G))$
  - we interpret disjunction using *max*:
    - $\mathrm{I}(F \vee G) = max(\mathrm{I}(F),\mathrm{I}(G))$
  - we interpret negation similarly to the fuzzy set calculation of complement
    - $\mathrm{I}(\neg F) = 1 - \mathrm{I}(F)$

Exercise

- Suppose, we have atoms {red, round} and
  - $\mathrm{I}(\text{red}) = 0.6$
  - $\mathrm{I}(\text{round}) = 1$

- We get
  - $\mathrm{I}(red \wedge round) = \min(\mathrm{I}(\text{red}), \mathrm{I}(\text{round})) = \min(0.6, 1) = 0.6$
  - $\mathrm{I}(\neg red) = 1.0 - \mathrm{I}(\text{red}) = 1.0 - 0.6 = 0.4$
  - $\mathrm{I}(red \vee \neg red) = \max(\mathrm{I}(\text{red}), \mathrm{I}(\neg red)) = \max(0.6, 1.0 - 0.6) = 0.6$
  - $\mathrm{I}(red \wedge \neg red) = \min(\mathrm{I}(\text{red}), \mathrm{I}(\neg red)) = \min(0.6, 1.0 - 0.6) = 0.4$
  - $\mathrm{I}(\text{red} \vee \neg round) = max(0.6, 1.0 - 1.0) = 0.6$

Fuzzy Set Theory and Fuzzy Logic

- The semantics for Fuzzy Logic operators can be defined in many different ways

- t-norms and s-norms generalize conjunction and disjunction

- A two-place operation t, resp. s on [0,1] is called t-norm, resp. s-norm if it holds:
  - 1 (for t) and 0 (for s) are neutral elements.
  - t and s are commutative.
  - t and s are associative.
  - t and s are monotone increasing:
    - Meaning: $x \leq x' \wedge y \leq y' \rightarrow t(x,y) \leq t(x',y')$

- Example: Multiplication of natural numbers has a neural element, is commutative, associative, and is monotonic increasing.

Fuzzy Set Theory and Fuzzy Logic

- Examples of *t*-norms and *s*-norms are
  - *min* is a *t*-norm and *max* is a *s*-norm
  - $alg_t(x,y) = x \cdot y$ is a *t*-norm and $alg_s(x,y) = x + y - (x \cdot y)$ is an *s*-norm.
  - $quo_t(x,y) = (xy) / (x + y - xy)$ is a *t*-norm and
    $quo_s(x,y) = (x + y - 2xy) / (1.0 - xy)$ is a s-norm.
- There are infinitely many possible t-norms and s-norms.
- Nevertheless, there is, for example, a smallest (non-trivial) t-norm and a largest (non-trivial) t-norm:
- Smallest *t*-norm: $t(x,y) = 1.0$ iff $x = 1.0$ and $y = 1.0$, else $t(x,y) = 0$
- Largest *t*-norm: $t(x,y) = \min(x,y)$

Applications of Fuzzy Logic

- Fuzzy logic is widely used for controlling dynamical systems.
  - Temperature control in air-conditioning systems (Mitsubishi, Sharp).
  - Stable control of car engines (Nissan).
  - Recognition of handwritten symbols (Sony).
  - Motor control of vacuum cleaners with recognition of surface conditions and degree of soiling (Matsushita).
  - Efficiency of elevator control (Fujitec, Hitachi, Toshiba).

- A recent book on applications is
  - Carter et al. (2021): *Fuzzy Logic: Recent Applications and Developments*, Springer Nature.

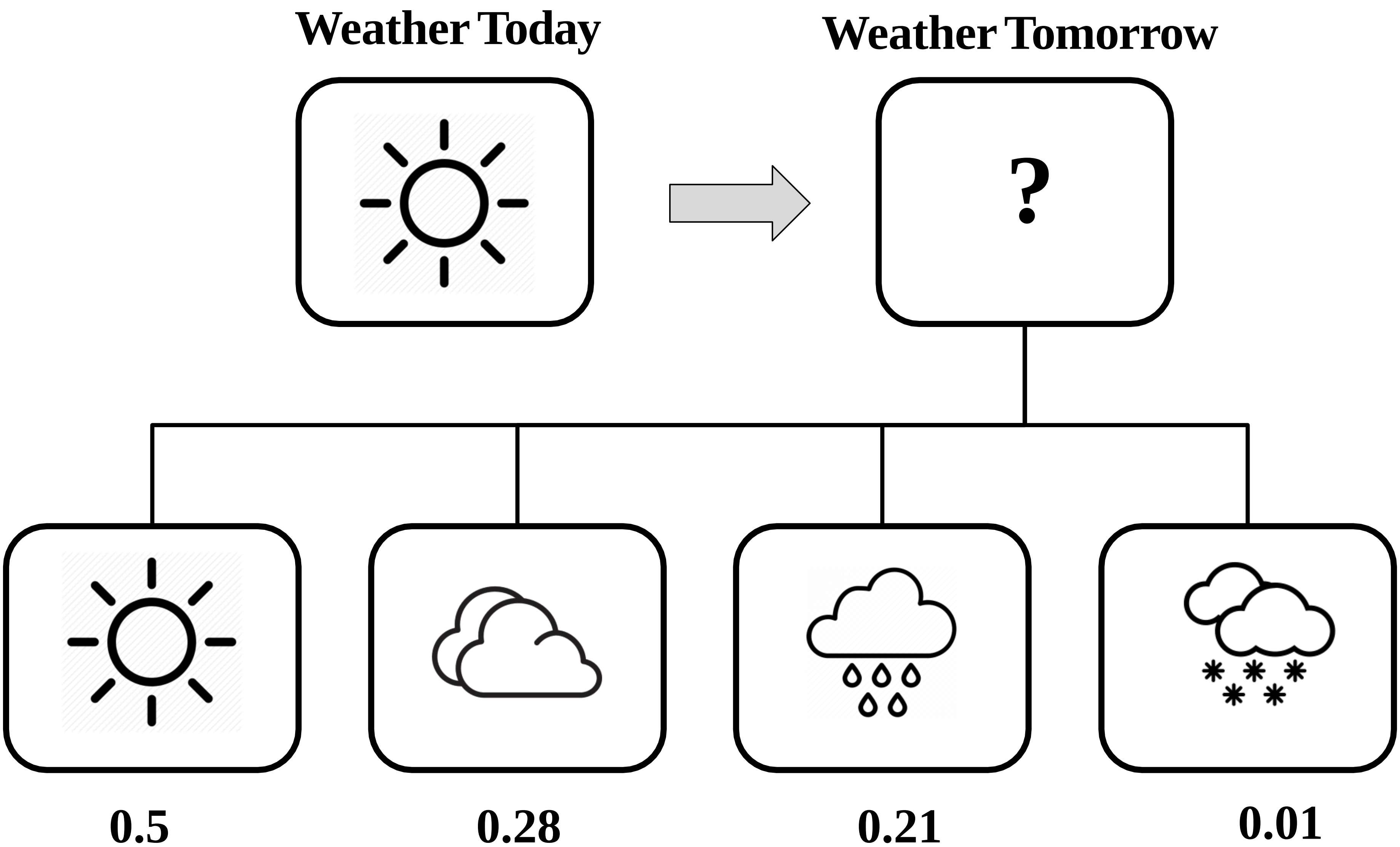
Summary

- Fuzzy Theory allows us to express vagueness
- Vague statements can be true to a certain degree
- This is accomplished by means of membership functions
- Fuzzy logics can be defined in many different ways
- However, usually we try to extend classical logics
  - For membership 0 and 1, behaviour like classical logic
  - In between, novel things can happen (e.g. $I(F \vee \neg F)$)

Probabilistic Logic: Introduction / Motivation

Modeling Uncertainty

Weather Example



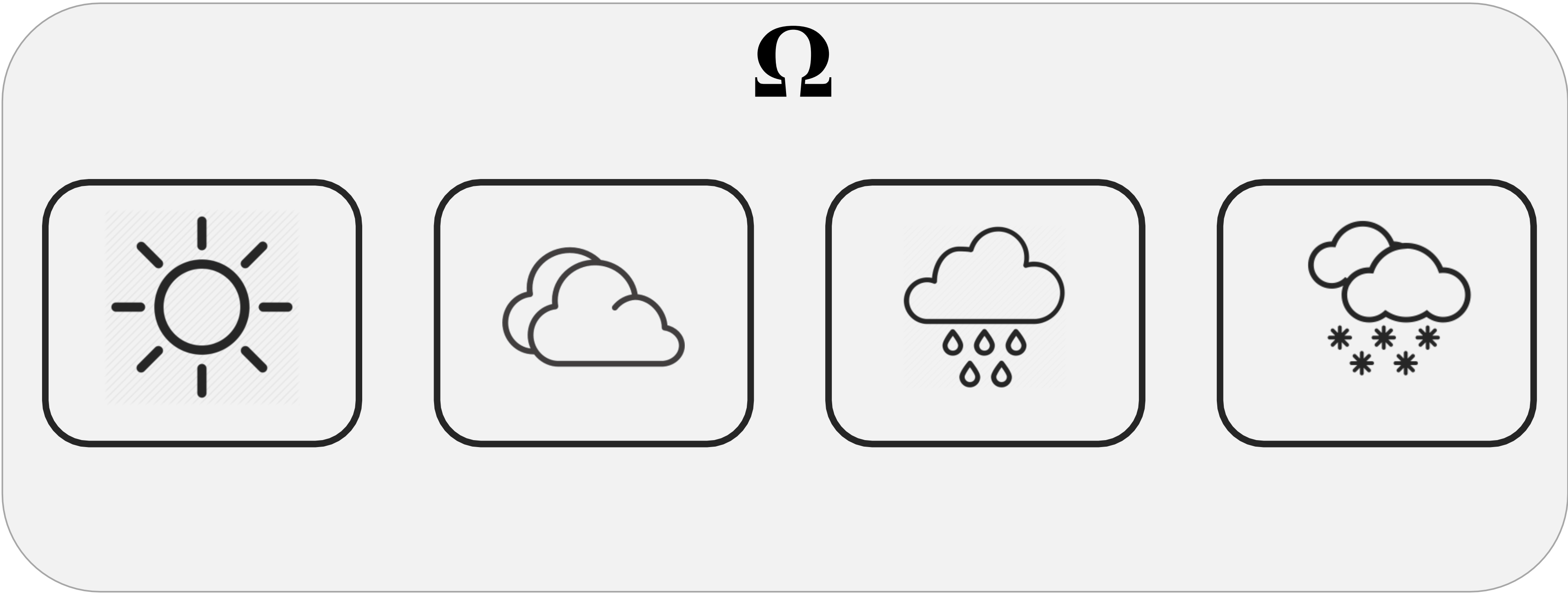
Probability Space

- A (simple finite) probability space is a triple $<\Omega,\Sigma,P>$
  - Ω is a finite non-empty set
  - Σ is the set of all subsets of Ω
  - $P: \Sigma \to \mathbf{R}$ is a '*probability measure*'

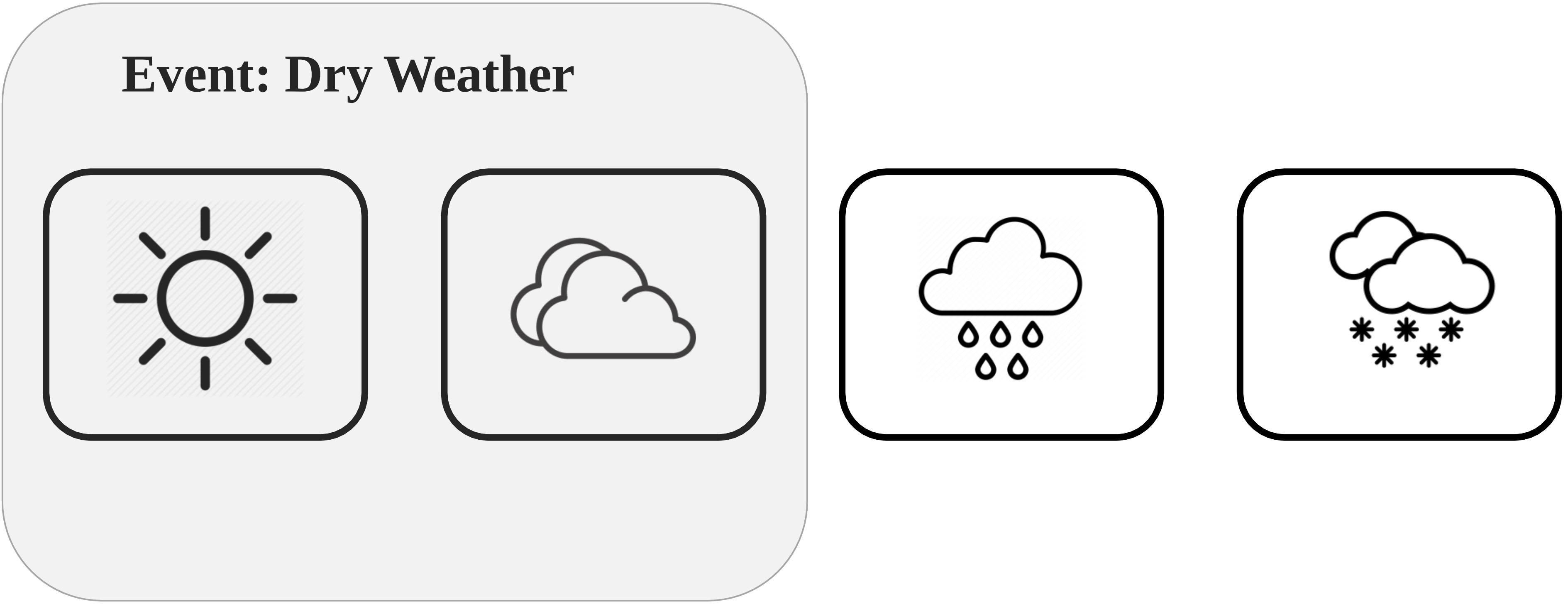
- Ω is the set of elementary events
  (in other fields also called «possible worlds" or "states")

- Σ is the set of events
  - Each event E in Σ is a subset of Ω
  - Intuitively, each event E corresponds to a set of possible worlds that satisfy some statement

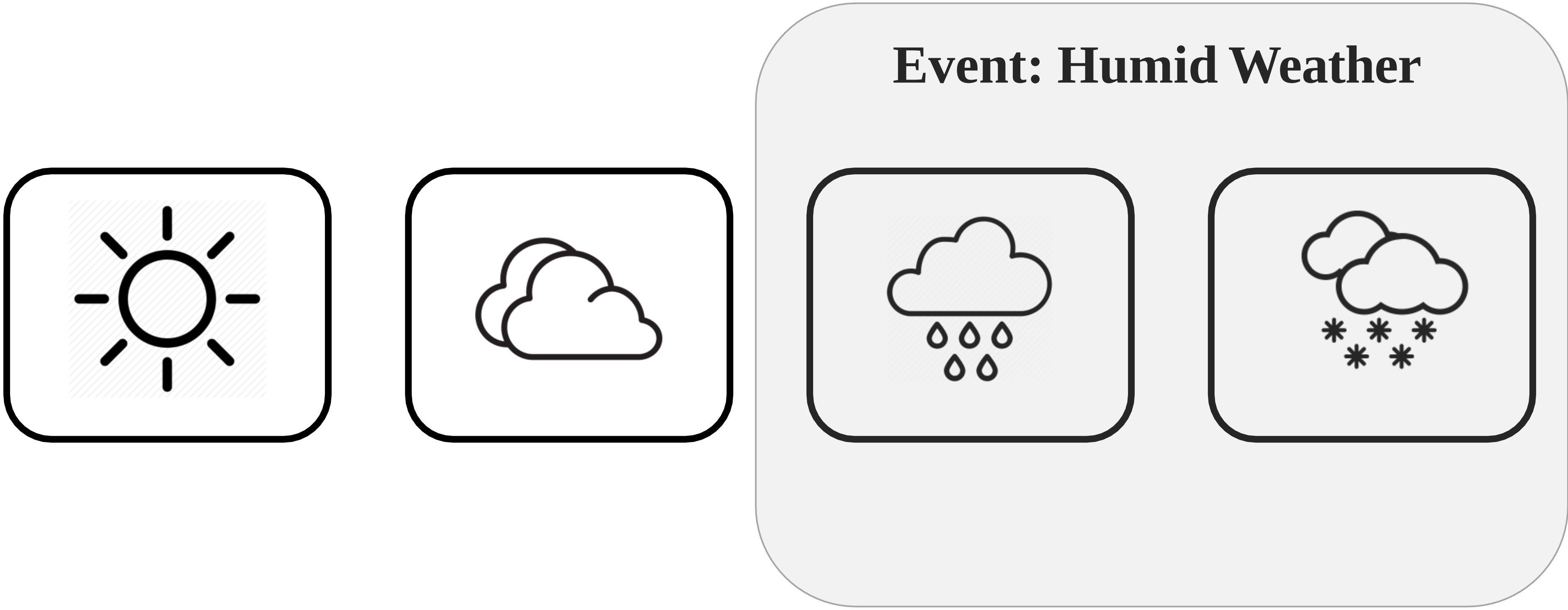
Possible Worlds



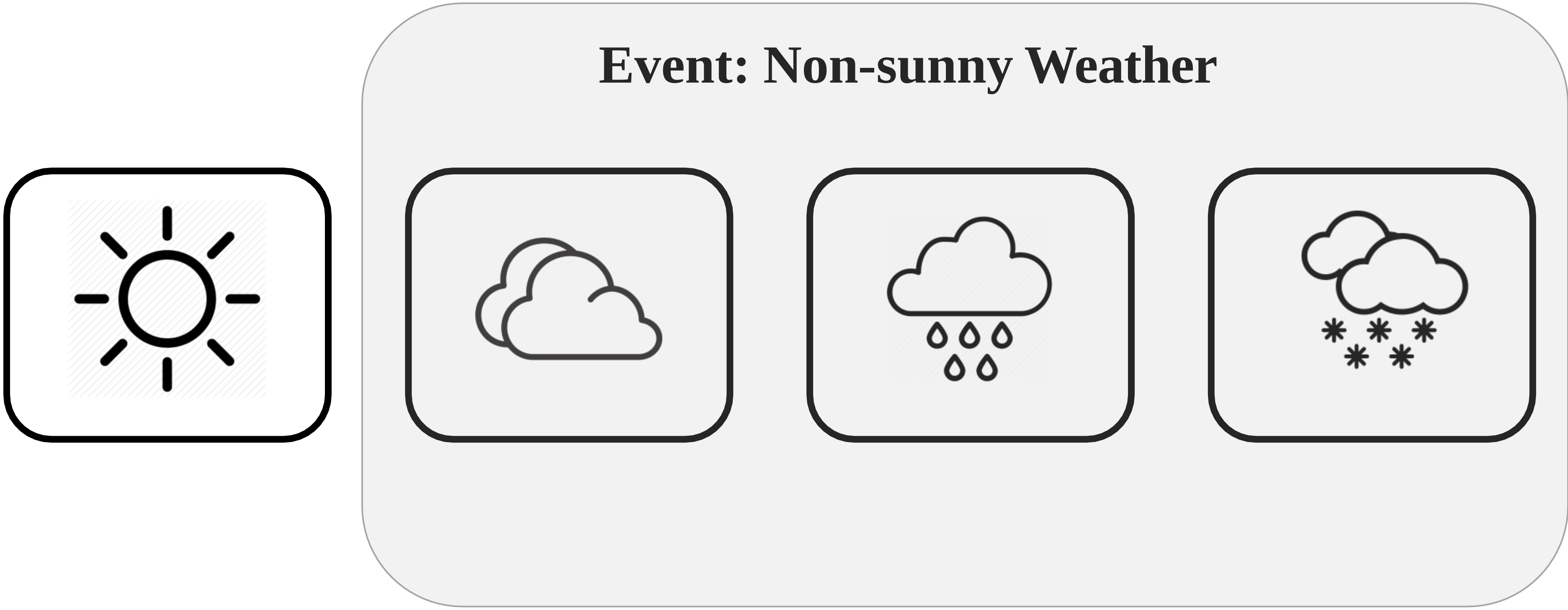
Events



Events



Events



Probabilistic Logic:
Probability Measures

Some more Formal Stuff

Probability Measures

- Probability measures assign probabilities to events
- They are characterized by Kolmogorov's axioms (finite version)

1. $P(E) \geq 0$ (Non-negativity)

2. $P(\Omega) = 1$ (Normalization)

3. If the events $E_1, E_2, \ldots, E_n$ are disjoint, then (Finite Additivity)

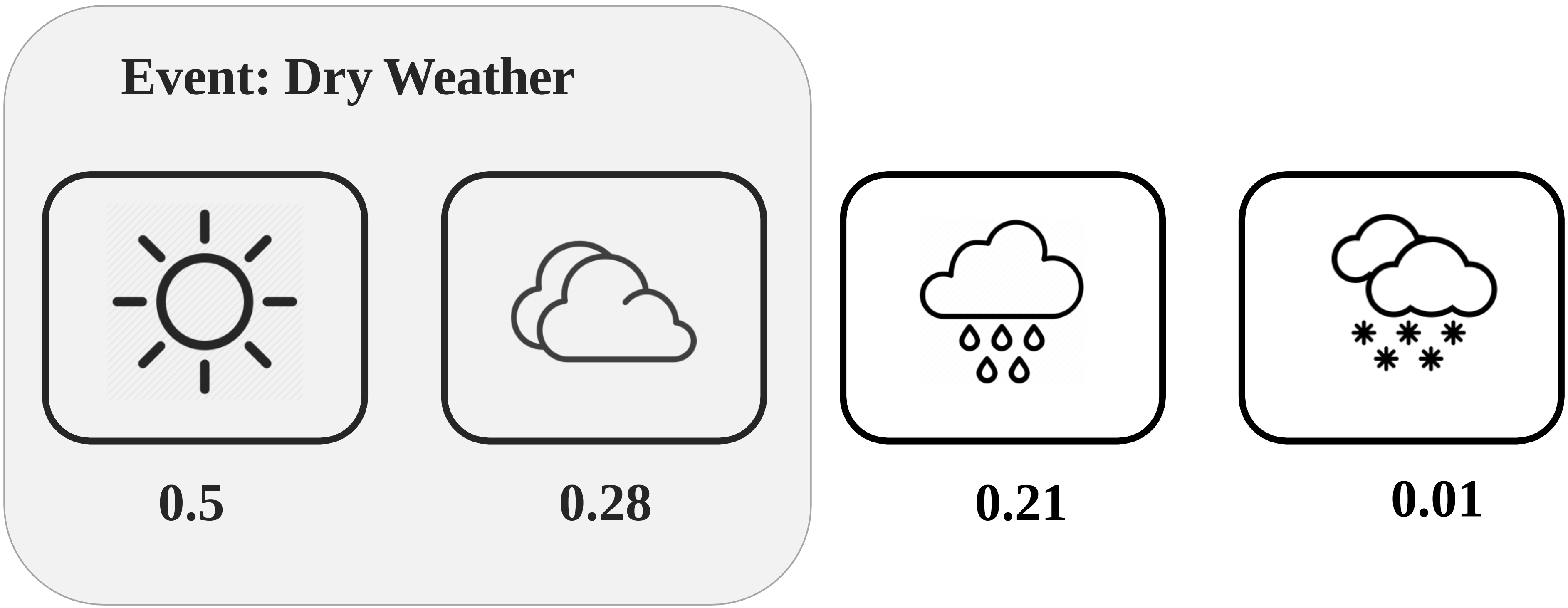
$$P(E_1 \cup E_2 \cup \ldots \cup E_n) = P(E_1) + P(E_2) + \ldots + P(E_n).$$

- The infinite version of Kolmogorov's axioms replaces the last axiom with:

The countable sequence of disjoint sets $E_1, E_2, \ldots$ satisfies:

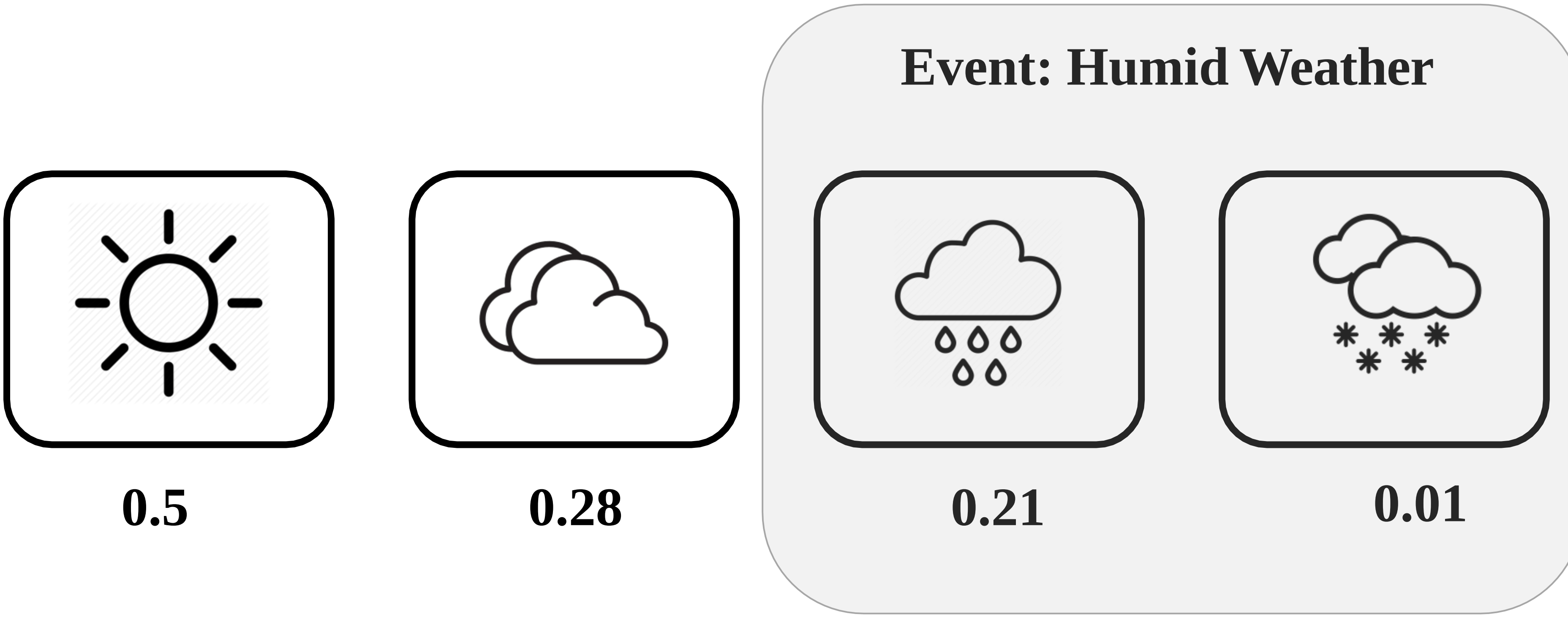
$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

Probability of Events



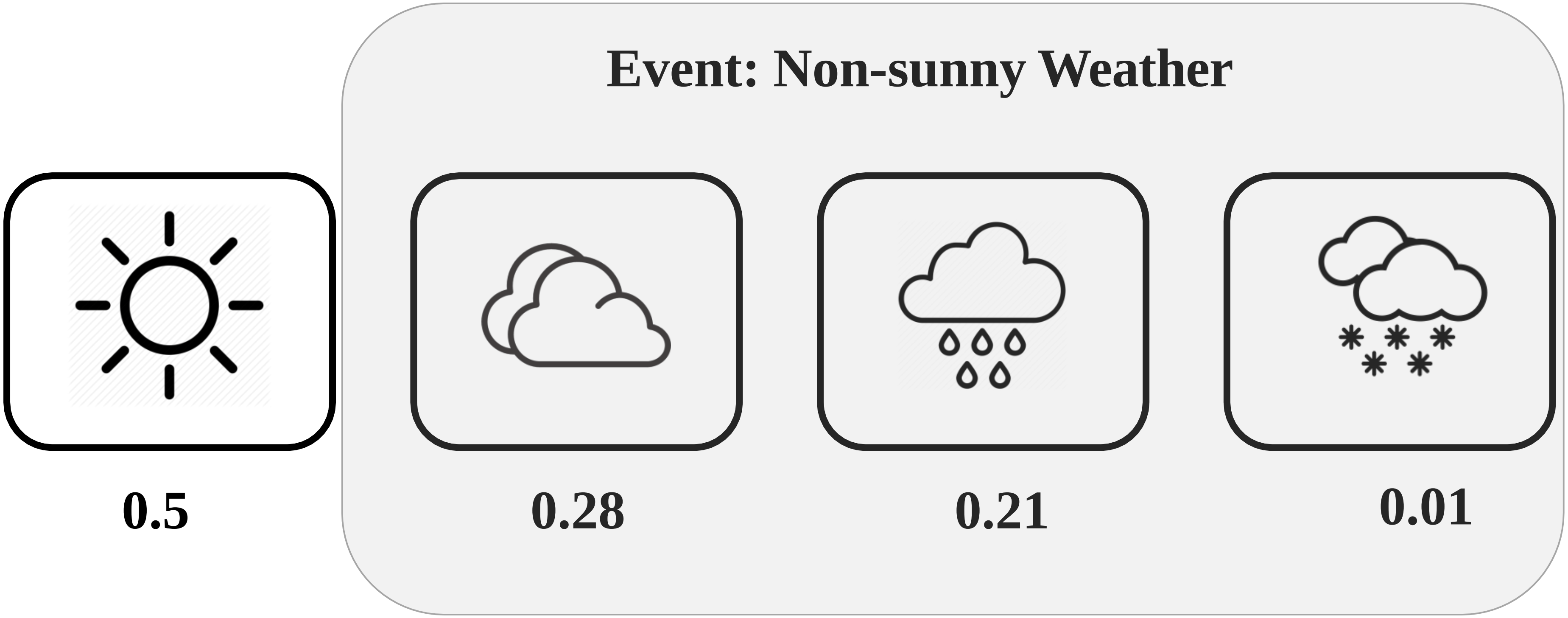
0.5 + 0.28 = 0.78

Probability of Events



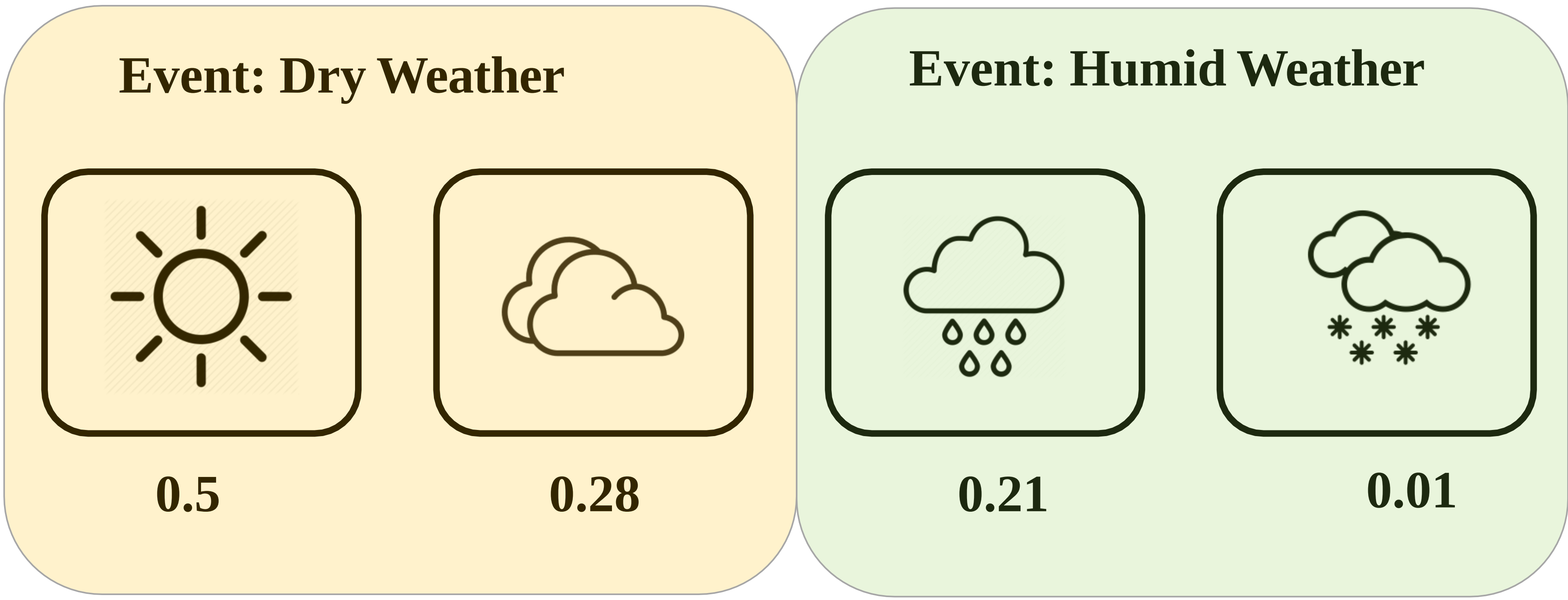
$$0.21 + 0.01 = 0.22$$

Probability of Events



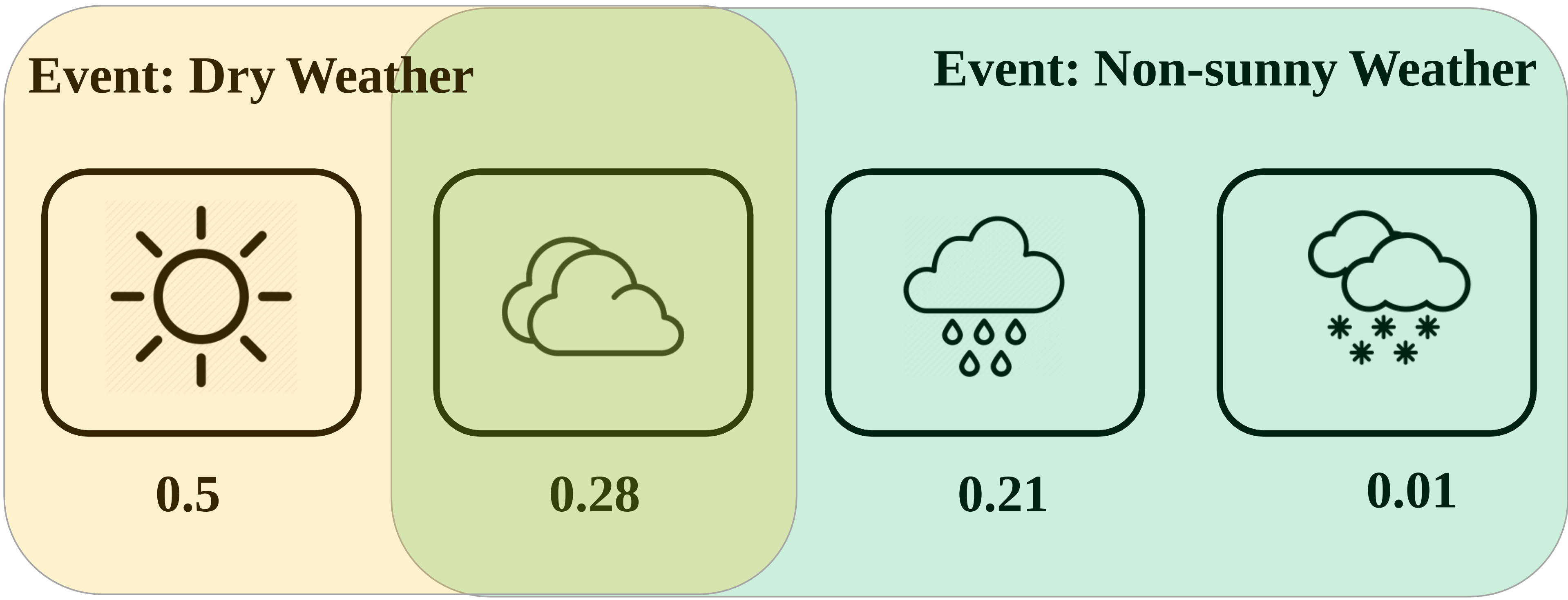
$$0.28 + 0.21 + 0.01 = 0.5$$

Additivity



P(Dry or Humid) = P(Dry) + P(Humid)

Additivity



P(Dry or Non-sunny) ≠ P(Dry) + P(Non-sunny)

Have a Nice Break and a Happy New Year !!!