符号学习简介

非经典逻辑与近似推理

(Press ? for help, n and p for next and previous slide)

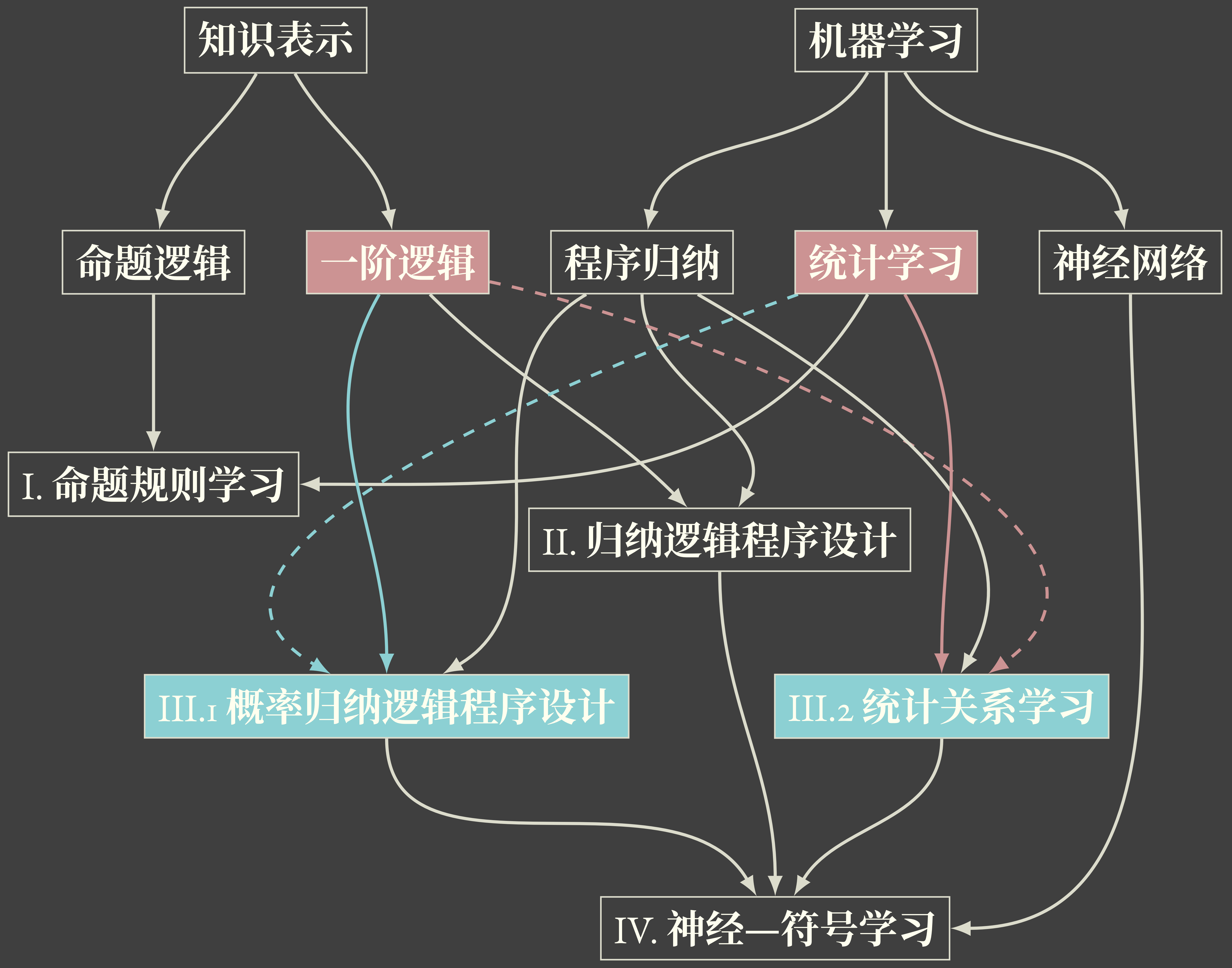
戴望州

南京大学智能科学与技术学院

2025年-秋季

https://daiwz.net

路径图



知识表示
机器学习
命题逻辑
一阶逻辑
程序归纳
统计学习
神经网络
I. 命题规则学习
II. 归纳逻辑程序设计
III.1 概率归纳逻辑程序设计
III.2 统计关系学习
IV. 神经–符号学习

符号学习简介・非经典逻辑与近似推理

1. **近似推理**
2. 多值逻辑
3. 模糊逻辑

近似推理

Łukasiewicz: “明年 12 月 21 日中午我将在华沙”

近似推理

“以后我会成为学者”

近似推理

“如果我会成为学者，大家一定会高兴的”

近似推理

"如果我会成为**比较厉害**的学者，
大家一定会**比较**高兴的"

近似推理

“如果扬声器有**比较严重**的交流声，
则滤波电容**多半失效**了”

近似推理

“如果扬声器有**不太严重**的交流声，
则滤波电容**可能容量不足**了”

近似推理

真值指派的结果不再只是$\{True, False\}$之中二选一。

在数学上应该如何刻画？

泛代数

设A为非空集，则:

1. A上的0元运算是A中的一个**元素**;
2. A上的n元运算（$n \geq 1$）是A中的一个n元**函数**$f : A^n \mapsto A$;

设Ω是非空集，N是非负整数集，$ar : T \mapsto N$是映射，则称(Ω, ar)为**型**（type, or signature）。有时把它简记为Ω。

设Ω为型，A为非空集。如果对每个$t \in \Omega$有一个$ar(t)$元函数$t_A : A^{ar(t)} \mapsto A$，则称$A$为$\Omega$型**泛代数**。当$t \in \Omega_n$时，$t_A$是$\Omega$泛代数$A$上的$n$元函数，其中0元运算即为$A$中的元素。

格

设 $\Omega = \{\top, \bot, \vee, \wedge\}$，$\Omega_0 = \{\top, \bot\}$，$\Omega_2 = \{\vee, \wedge\}$，则任一有界格（Lattice）$L$ 都是一个 Ω 型泛代数，$\top_L$ 和 $\bot_L$ 分别是 L 上的最大、最小元，$\vee_L$ 和 $\wedge_L$ 分别是 L 中的上、下确界运算。

布尔代数

布尔代数具有以下性质:

1. 是一个完备格:
    - 有偏序、最大元、最小元;
    - 函数为上确界运算、下确界运算;
    - 任意子集均存在上下确界;
2. 有补运算 $\neg$;
3. 是分配格(即满足 de Morgan 律);

泛代数同态

设A和B都是具有同一型Ω的泛代数，$\varphi : A \mapsto B$是映射，若

1. 对每个0元函数$t \in \Omega_0$，对任意t_A存在t_B令$\varphi(t_A) = t_B$
2. 对每个n元函数$t \in \Omega_n, n \geq 1$，对任意A中n个元素$a_1, \ldots, a_n$有
   $\varphi(t_A(a_1, a_2, \ldots, a_n)) = t_B(\varphi(a_1), \varphi(a_2), \ldots, \varphi(a_n))$

则称φ为从A到B的一个Ω型<u>同态</u>。如果φ即单又满，则称其为Ω型<u>同构</u>。

真值指派

1. 逻辑语言是由 $S = \{p_1, p_2, \ldots\}$ 生成的 $\Omega = \{\neg, \to\}$ 型**自由代数**。
2. 真值指派 v 是将以上自由代数映射至布尔代数的 Ω 型同态。

近似推理 vs 经典逻辑

否定排中律：$v(p \lor \neg p) \neq \top$

近似推理 vs 经典逻辑

真值指派所使用的格（赋值格）是布尔代数的**扩张**

- 即布尔代数应当是近似推理方法赋值格的一个**子代数**

用于近似推理的非经典逻辑

- 模态逻辑（Modal Logic）：通过模态词来限定真值
- 多值逻辑（Many-valued Logic）：$L = \{0, I_1, I_2, \ldots, 1\}$（赋值格为有限或可数个真值）
- 模糊逻辑（Fuzzy Logic）：$L = [0, 1]$（赋值格为连续值）
- 概率逻辑（Probabilistic Logic）：使用概率论作为语义基础
- ……

符号学习简介・非经典逻辑与近似推理

1. 近似推理
2. **多值逻辑**
3. 模糊逻辑

多值逻辑

- 7是素数
- 9是素数
- 11是水仙花

- $\top$：真、true、T、1
- $\bot$：假、false、F、0
- I：不确定、indeterminate、I、1/2

Łukasiewicz 的三值系统 L_3

p	$\neg p$
T	F
I	I
F	T

Łukasiewicz 的三值系统 L_3

第一列为 p 的真值，第一行为 q 的真值:

$p \vee q$	T	I	F
T	T	T	T
I	T	I	I
F	T	I	F

Łukasiewicz 的三值系统 L_3

第一列为p的真值，第一行为q的真值:

$p \wedge q$	T	I	F
T	T	I	F
I	I	I	F
F	F	F	F

Łukasiewicz 的三值系统 L_3

第一列为 p 的真值，第一行为 q 的真值：

$p \to q$	T	I	F
T	T	I	F
I	T	T	I
F	T	T	T

Łukasiewicz 的三值系统 L_3

1. 是二值布尔逻辑的直接扩充
2. MP规则成立、de Morgan律成立
3. 连接词的语义：
   - $p \vee q = (p \to q) \to q$
   - 实质蕴涵

$$p \to q = \begin{cases} \neg p \vee q, & (p, q) \neq (I, I) \\ T, & (p, q) = (I, I) \end{cases}$$

L_3 中的模态词

- 必然$\Box$:

p	$\Box p$
T	T
I	F
F	F

- 或然$\Diamond$:

p	$\Diamond p$
T	T
I	T
F	F

L_3 中的定理

- $p \to (q \to p)$
- $p \vee \neg p$
- $\neg(p \to \neg p) \vee \neg(\neg p \to p)$

多值逻辑中的定理

二值逻辑中许多定理（重言式）不再是多值逻辑里的重言式。于是我们称：

- 真值恒为T的成为**重言式**
- 真值恒大于F的成为**准重言式**

Kleene 的三值系统 K_3

仅蕴涵算子与 L_3 不同:

$p \to q$	T	I	F
T	T	I	F
I	T	I	I
F	T	T	T

K_3中的准重言式

K_3中有$p \vee q = \neg p \to q$以及$p \to q = \neg p \vee q$，因此K_3可看作仅含有连接词$\neg$与$\to$。

定理：K_3中的准重言式与二值逻辑中的的重言式一致。

Łukasiewicz 的 n 值系统 L_n

将赋值格由 $\{T, I, F\}$ 推广为含有 n 个元素的线性序集：

$$L = \{0, \alpha_1, \alpha_2, \ldots, \alpha_{n-2}, 1\},$$

其中 0 表示假，1 表示真，且

$$\alpha_o = 0 < \alpha_1 < \ldots < \alpha_{n-2} < \alpha_{n-1} = 1$$

一般定义它们为 $[0, 1]$ 区间的 n 等分点，这时有：

- $\neg\alpha = 1 - \alpha$
- $\alpha \wedge \beta = \min\{\alpha, \beta\}$
- $\alpha \vee \beta = \max\{\alpha, \beta\}$
- $\alpha \rightarrow \beta = \min\{\neg\alpha + \beta, 1\}$

L_n 的子代数

L_n是$(\neg, \vee, \rightarrow)$型代数，设$M$是$L_n$的非空子集，如果$M$关于运算$(\neg, \vee, \rightarrow)$封闭，则称$M$是$L_n$的子代数。

L_n是L_{2n-1}的子代数。二值逻辑代数C_2是$L_k, k > 2$的最小子代数。

定理：设L_n是L_m的子代数，$T(L_n)$代表L_n中所有的重言式，那么有：

$$T(L_m) \in T(L_n)$$

L_n 的子代数

如果 L_n 不是 L_m 的子代数，则重言式之间没有以上关系。比如令 $m=4, n=3$:

- $A = ((p \to \neg p) \land (\neg p \to p)) \to (\neg p \lor p)$,
- $B = \neg p \lor ((p \to \neg p) \to (\neg p \to p))$

多值逻辑上的重言式

定义：设A是S生成的自由代数中$F(S)$的公式，L是某种线性序的n值系统，$\alpha \in L$。如果对每个赋值$v : F(S) \mapsto L$恒有$v(A) \geq \alpha$，则称A为α重言式。

其它赋值格上的算子

一般而言，$\alpha, \beta \in [0,1]$时:

- $\neg\alpha = 1 - \alpha$
- $\alpha \vee \beta = \max\{\alpha, \beta\}$
- $\alpha \wedge \beta = \min\{\alpha, \beta\}$
- 但是在多值逻辑赋值格上对蕴涵$\alpha \to \beta$的定义则多种多样

赋值格上的蕴涵算子

多值逻辑系统赋值格上各家对$\alpha \to \beta$的定义:

算子	定义
Zadeh算子	$\neg\alpha \vee (\alpha \wedge \beta)$
Łukasiewicz算子	$(\neg\alpha + \beta) \wedge 1$
Mamdani算子	$\alpha \wedge \beta$
Gaines-Rescher算子	$1 \text{ if } \alpha \leq \beta; 0 \text{ if } \alpha > \beta$
Reichenbach算子	$\neg\alpha + \alpha \cdot \beta$
Gödel算子	$1 \text{ if } \alpha \leq \beta, \beta \text{ if } \alpha > \beta$
Goguen算子	$1 \text{ if } \alpha = 0, \frac{\beta}{\alpha} \wedge 1 \text{ if } \alpha > 0$
Yager算子	β^{α}
Kleene-Dienes算子	$\neg\alpha \vee \beta$

Dubois-Prade条件

直觉上，多值逻辑中蕴涵算子应当满足的10个条件：

1. $\alpha \leq \gamma$时，$\alpha \to \beta \geq \gamma \to \beta$
2. $\beta \leq \gamma$时，$\alpha \to \beta \leq \alpha \to \gamma$
3. $0 \to \beta = 1$
4. $1 \to \beta = \beta$
5. $\alpha \to \beta \geq \beta$
6. $\alpha \to \alpha = 1$
7. $\alpha \to (\beta \to \gamma) = \beta \to (\alpha \to \gamma)$
8. $\alpha \to \beta = 1$当且仅当$\alpha \leq \beta$
9. $\alpha \to \beta = \neg\beta \to \neg\alpha$
10. $\alpha \to \beta$关于α, β连续

符号学习简介・非经典逻辑与近似推理

1. 近似推理
2. 多值逻辑
3. **模糊逻辑**

模糊逻辑（Fuzzy Logic）

1. 将命题的真值指派扩充为 $[0, 1]$ 区间的连续值；
2. 利用模糊算子（Fuzzy operators）对公式进行推理。

模糊推理的基本思想

经典逻辑中我们有肯定前件（MP）和演绎定理：

$$\begin{array}{r} A \to B \\ A \\ \hline B \end{array}$$

模糊推理的基本思想

如果出现这样的情况：

$$\frac{\begin{array}{r} A \to B \\ A' \end{array}}{B'}$$

它还成立吗？

模糊推理的基本思想

模糊逻辑中，确实会出现这样的情况：

$$\frac{\begin{array}{r} R(A,B) \\ A' \end{array}}{B'}$$

其中 A' 和 B' 分别是对 A 与 B 在某种“程度”上的描述。

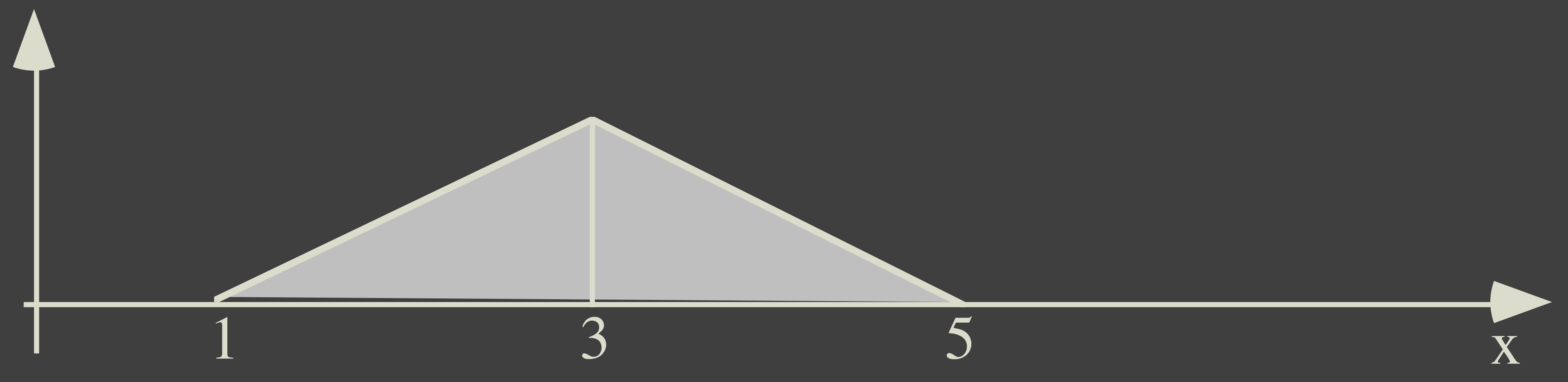
模糊集（Fuzzy Set）

定义：给定一个论域U，那么从U到单位区间$[0,1]$的一个映射$\mu_A : U \mapsto [0,1]$称为U上的一个模糊集，或U的一个模糊子集。

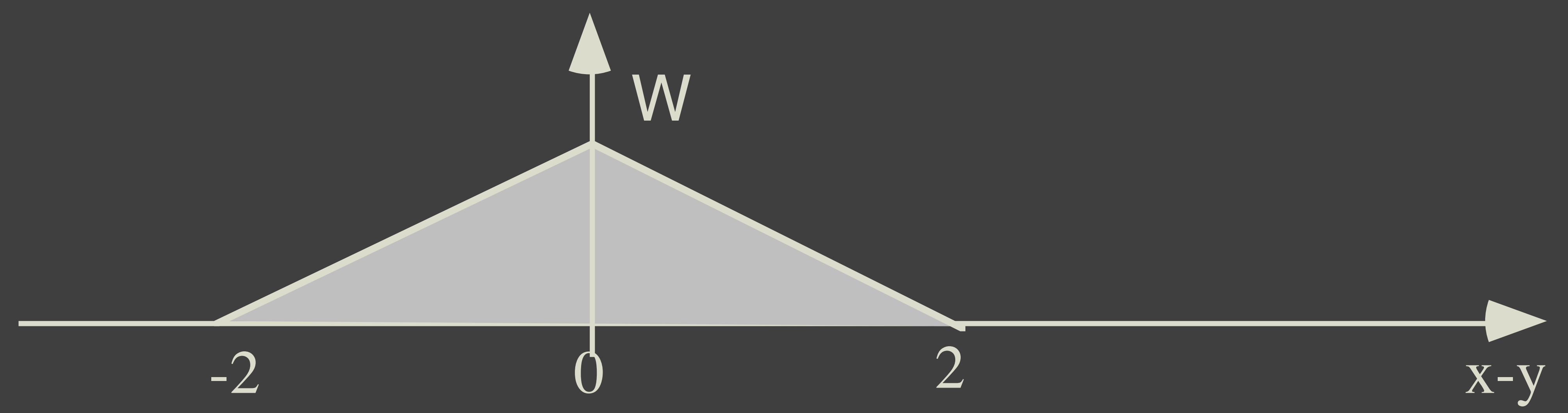
- 模糊集可以记为A
- 映射（函数）$\mu_A(\cdot)$或简记为$A(\cdot)$叫做模糊集A隶属函数
- 对于每个$x \in U$，$\mu_A(x)$叫做元素x对模糊集A的隶属度。

模糊化

$$\frac{\begin{array}{r} X \text{ 离 } Y \text{ 很近} \\ X \text{ 离 } 3 \text{ 很近} \end{array}}{Y \text{ 离 } 3 \text{ 很近}}$$



X离3很近



X-Y离0很近

模糊推理的CRI方法

$$\{A \to B\} \cup \{A'\} \vdash_{\text{fuzzy}} B'$$

对于上面的问题，模糊逻辑中采用合成推理（compositional rule of inference, CRI）方法：

1. 将命题用模糊集表示（模糊化），例如$A, A' \in \mathfrak{F}(X)$，$B, B' \in \mathfrak{F}(Y)$
2. 把蕴涵式$A \to B$转化为一个$X \times Y$上的模糊关系$R : X \times Y \mapsto [0,1]$，例如Zadeh蕴涵算子

$$R(x,y) = A'(x) \vee (A(x) \wedge B(y))$$

3. 把给定的$A'(x)$（即输入）与上一步中的模糊集做合成即得到输出$B' = A' \circ R$。Zadeh的合成算法是：

$$\begin{aligned} B &= \sup_{x \in X, y \in Y} [A'(x) \wedge R(x,y))] \\ &= \sup_{x \in X, y \in Y} \{A'(x) \wedge [A'(x) \vee (A(x) \wedge B(y))]\} \end{aligned}$$

T-NORM

在模糊推理中除了蕴涵以外，还有许多对合取（$\wedge$）的定义方式。它们统称为“三角范数”（triangular-norm, t-norm），是一种$[0,1]\times[0,1]\mapsto[0,1]$的映射，需满足如下条件：

1. 有界：$T(0,0)=0, T(a,1)=T(1,a)=a$
2. 单调：$(a\leq c\wedge b\leq d)\to T(a,b)\leq T(c,d)$
3. 对称：$T(a,b)=T(b,a)$
4. 结合律：$T(a,T(b,c))=T(T(a,b),c)$

例如：

- Łukasiewicz T范数：$T(x,y)=\min\{x,y\}$
- 概率T范数：$T(x,y)=xy$
    - 若定义$\neg x=1-x$，则$x\vee y$就是noisy-or
- 有界T范数：$T(x,y)=\max\{0,x+y-1\}$

到底什么是模糊推理？

模糊推理的数学本质只是

在模糊集上定义的某种映射

连续化逻辑的根本问题

准重言式

若S为无限集，$F(S)$为由它生成的$(\neg,\vee,\rightarrow)$型自由代数，$[0,1]$上的运算$\rightarrow$由某个连续蕴涵算子确定，Σ是若干赋值$v:F(S)\mapsto[0,1]$的集合，$A,B\in F(S)$，$\alpha\in(o,1]$，定义：

- 称对每个真值指派v，$v(A)\geq\alpha, v(A\rightarrow B)\geq\alpha$推得$v(B)\geq\alpha$的规则为$\Sigma$中的$\alpha$-MP规则（严格大于则称为$\alpha^+$）。
- 称对每个真值指派v，$v(A\rightarrow B)\geq\alpha, v(B\rightarrow C)\geq\alpha$推得$v(A\rightarrow C)\geq\alpha$的规则为$\Sigma$中的$\alpha$-HS规则（严格大于则称为$\alpha^+$）。

连续逻辑与经典逻辑不兼容

定理：设$R(x,y)$关于y不减，$v(A\to B)$由$R(v(A),v(B))$定义。如果有$A,B\in F(S)$以及$v\in\Sigma$使$v(A)=\frac{1}{2},v(B)=0$，则以下两个条件不能同时成立：

1. $\frac{1}{2}^{+}$-MP规则与$\frac{1}{2}^{+}$-HS规则成立
2. 以下三条经典逻辑公理均为$\frac{1}{2}^{+}$重言式：
   - （L1）$A\to(B\to A)$
   - （L2）$(A\to(B\to C))\to((A\to B)\to(A\to C))$
   - （L3）$(\neg B\to\neg A)\to(A\to B)$

大部分情况下，如果要令$\frac{1}{2}^{+}$-MP规则与$\frac{1}{2}^{+}$-HS规则成立，则必须放弃公理L2

- 显然，因为随着蕴涵关系的传递，真值会衰减
- 因此多值/连续逻辑不再有经典逻辑里“蕴涵”对于“有效性”的保障

小结

非经典逻辑部分小结

1. 逻辑和赋值可看作泛代数之间的同态映射
2. 非经典逻辑是对布尔代数的延拓
   - 基本无法做到无损扩充
3. 公式的真值也被连续化后，“entailment”不再能保证“validity”