

人工智能导论 2018 期末考试答案

一、(基本定义)(10 分)

什么是智能 Agent? 智能 Agent 有哪几种, 简述其原理。

• **Agent:** 感知和行为的实体; 或者, 可以被视为感知和行动的一个。基本上任何对象都有资格; 关键是对对象实现代理功能的方式。(注: 有些作者将该术语限制为代表人工作的程序, 或者限制可能导致部分或全部代码在网络上的其他计算机上运行的程序, 如在移动 Agent 中)。

• **Agent 功能:** 一种函数, 用于指定代理响应每个可能的感知序列的操作。

• **Agent 程序:** 该程序结合机器体系结构实现代理功能。在我们简单的设计中, 程序在每次调用时都会获得一个新的感知并返回一个操作。

• **合理性:** 鉴于迄今为止的知识, 选择能最大化其预期效用的行动的代理人的财产。

• **自主性:** Agent 的性质, 其行为取决于他们自己的经验, 而不仅仅是他们最初的规划。

• **反射 Agent:** 其行为仅取决于当前感知的代理。

• **基于模型的 Agent:** 一种 Agent, 其行为直接来自当前世界状态的内部模型, 并随时间更新。

• **基于目标的 Agent:** 选择其认为会实现明确表示的目标的 Agent。

• **基于效用的 Agent:** 选择它相信的操作的代理将最大化结果状态的预期效用。

• **学习 Agent:** 根据其经验, 其行为随时间推移而改善的 Agent。

二、(约束满足问题)(30 分)

三个本地人和三个外地人在河的一边, 还有一条能载一个人或者两个人的船。找到一个办法能使所有的人都渡到河的另一岸, 要求在任何地方外地人数都不能多于本地人数 (可以有外地人没有本地人)。

1) 请给出问题的形式化定义, 画出该问题的完全状态空间图。(12)

2) 分别用深度优先搜索、宽度优先搜索、启发式搜索算法求解该问题, 比较上述三种算法的完备性和最优性。(18)

答: 形式化定义如下:

状态: 用一个三元组 $s_k = (m, c, b)$ 来表示河岸上的状态, 其中 $m, c \in \{0, 3\}$ 分别代表某一岸上本地人与外地人的数目, $b=1$ 表示船在左岸, $b=0$ 则表示船在右岸。

初始状态: $S_0 = (3, 3, 1)$

行动: 把船从左岸划向右岸定义为 P_{ij} 操作。其中, 第一个下标 i 表示船载的本地人数, 第二个下

标 j 表示船载的外地人数; 同理, 从右岸将船划回左岸称之为 Q_{ij} 操作。则共有 10 种操作, 操作集

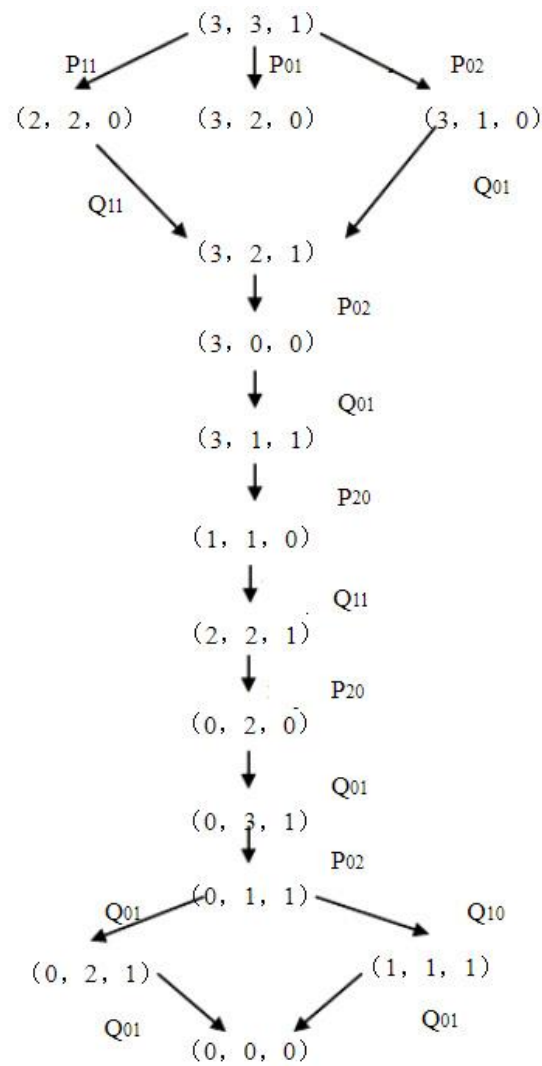
为:

$$F = \{P_{01}, P_{10}, P_{11}, P_{02}, P_{20}, Q_{01}, Q_{10}, Q_{11}, Q_{02}, Q_{20}\}$$

目标测试: 确定给定状态是不是目标状态 $(0, 0, 0)$

路径消耗: 划船次数

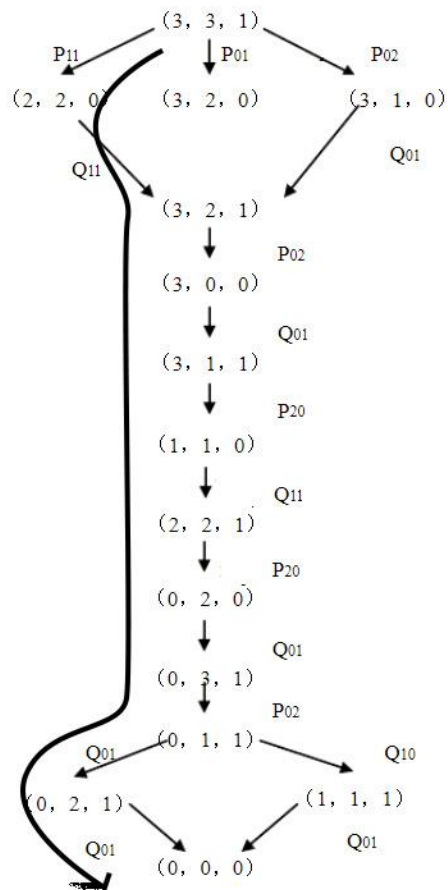
状态空间图：



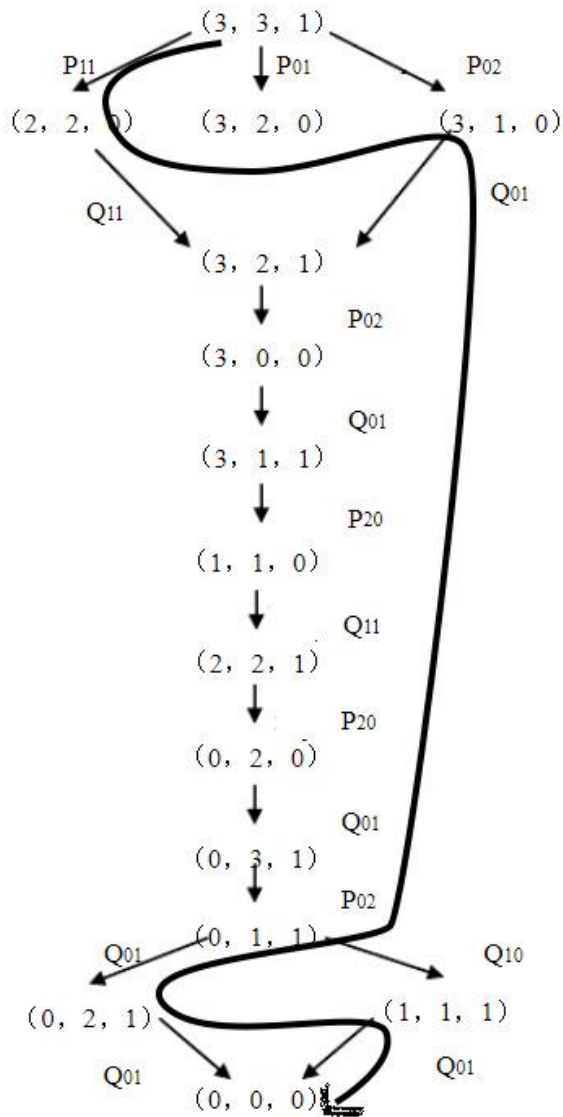
- 1) 分别用深度优先搜索、宽度优先搜索、启发式搜索算法求解该问题，比较上述三种算法的完备性和最优性。（18）

答：

DFS 搜索过程如下所示，dfs 在有限状态空间是完备、不是最优的。在该题是完备、最优的。



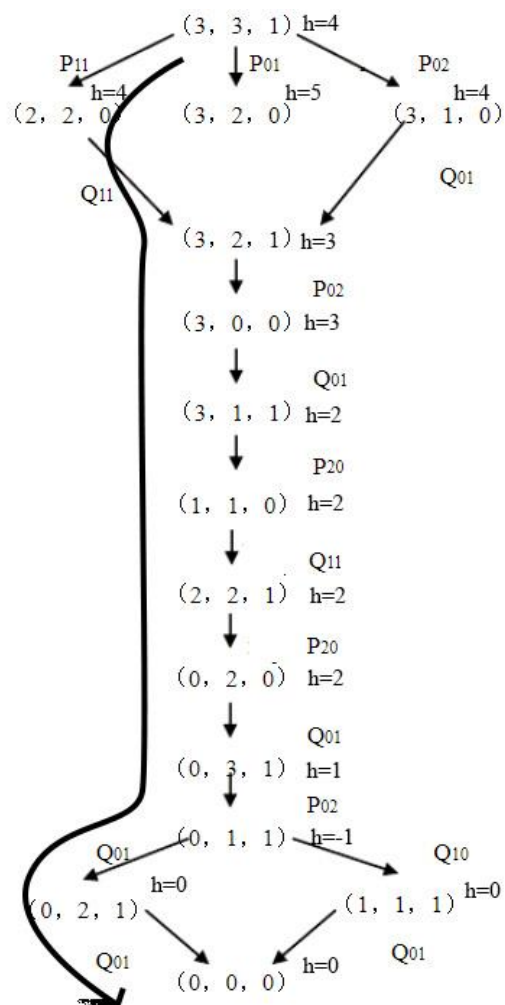
BFS 搜索过程如下，对于该题，路径代价是基于结点深度的非递减函数，是完备、最优的。



启发式：

定义启发函数 $h(s) = m + c - 2 * b$

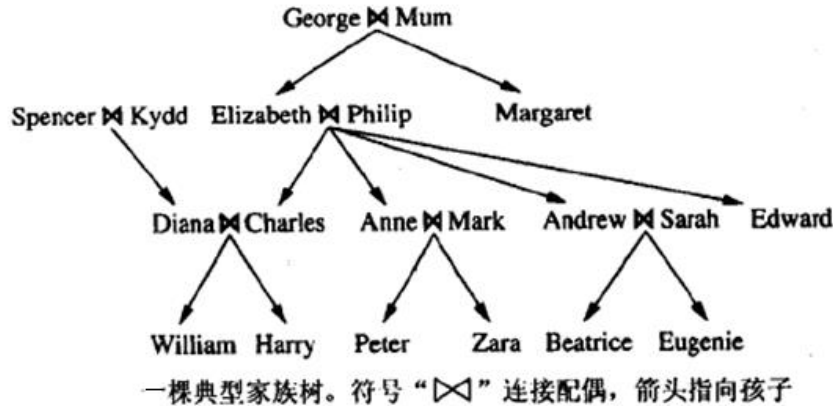
采用贪婪最佳优先搜索路径如下，一般来说贪婪最佳优先搜索不是完备、不是最优的，但在该题中完备最优。



三、（一阶逻辑）（20）

给出描述谓词 GrandChild（孙子女）、GreatGrandparent（曾祖父母）、Brother（兄弟）、Sister（姐妹）、Daughter（女儿）、Son（儿子）、Aunt（姑/姨）、Uncle（叔/舅）、BrotherInLaw（姐夫/妹夫）、SisterInLaw（兄嫂/弟妹）和 FirstCousin（第一代姑表亲）的公理。找到隔了 n 代的第 m 代姑表亲的合适定义，并用一阶逻辑写出该定义。

写出图中所示的家庭树的基本事实。采用适当的逻辑推理系统，把你已经写出的所有语句 TELL 系统，并 ASK 系统：谁是 Elizabeth 的孙子女，Diana 的姐夫/妹夫和 Zara 的曾祖父母？



【答案】

公理：

$Grandchild(c, a) \Leftrightarrow \exists b \ Child(c, b) \wedge Child(b, a)$
 $Greatgrandparent(a, d) \Leftrightarrow \exists b, c \ Child(d, c) \wedge Child(c, b) \wedge Child(b, a)$
 $Ancestor(a, x) \Leftrightarrow Child(x, a) \vee \exists b \ Child(b, a) \wedge Ancestor(b, x)$
 $Brother(x, y) \Leftrightarrow Male(x) \wedge Sibling(x, y)$
 $Sister(x, y) \Leftrightarrow Female(x) \wedge Sibling(x, y)$
 $Daughter(d, p) \Leftrightarrow Female(d) \wedge Child(d, p)$
 $Son(s, p) \Leftrightarrow Male(s) \wedge Child(s, p)$
 $FirstCousin(c, d) \Leftrightarrow \exists p_1, p_2 \ Child(c, p_1) \wedge Child(d, p_2) \wedge Sibling(p_1, p_2)$
 $BrotherInLaw(b, x) \Leftrightarrow \exists m \ Spouse(x, m) \wedge Brother(b, m)$
 $SisterInLaw(s, x) \Leftrightarrow \exists m \ Spouse(x, m) \wedge Sister(s, m)$
 $Aunt(a, c) \Leftrightarrow \exists p \ Child(c, p) \wedge [Sister(a, p) \vee SisterInLaw(a, p)]$
 $Uncle(u, c) \Leftrightarrow \exists p \ Child(c, p) \wedge [Brother(a, p) \vee BrotherInLaw(a, p)]$

定义每个人到其最近公共祖先的距离为 distance，例如定义 Distance(c,a)如下：

$Distance(c, c) = 0$
 $Child(c, b) \wedge Distance(b, a) = k \Rightarrow Distance(c, a) = k + 1.$

因此，到 grandparent 距离为 2，到 great-great-grandparent 距离为 4。

$MthCousinNTimesRemoved(c, d, m, n) \Leftrightarrow$
 $\exists a \ Distance(c, a) = m + 1 \wedge Distance(d, a) = m + n + 1.$

基本事实：

- 1、每个箭头代表两个孩子的例子。例如 Child(William,Diana) and Child(William,Charles)。
- 2、每个名字代表性别命题。例如 Male(William) or Female(Diana)。
- 3、 \bowtie 表示配偶命题。例如 Spouse(Charles,Diana)

下面给出几个 TELL 例：

TELL(KB,Female(Diana))

TELL(KB,Male(William))

.....

ASKVARS(KB,GrandChild(Elizabeth))

ASKVARS(KB,BrotherInLaw(Diana))

ASKVARS(KB,GreatGrandParent(Zara))

四、(决策树) (15 分)

决策树是一种混合算法，它综合了多种不同的创建树的方法，并支持多个分析任务，包括回归、分类以及关联。它也是机器学习中常用的算法之一。请依据对决策树的理解，回答以下问题。

- (1) 请用简短的语句描述熵和信息增益的定义。(5 分)
- (2) 下表为是否适合打垒球的决策表，请用决策树算法画出决策树。(6 分)
- (3) 请预测 $E = \{\text{天气}=\text{晴}, \text{温度}=\text{适中}, \text{湿度}=\text{正常}, \text{风速}=\text{弱}\}$ 的场合，是否合适中打垒球。(4 分)

天气	温度	湿度	风速	活动	天气	温度	湿度	风速	活动
晴	炎热	高	弱	取消	晴	适中	高	弱	取消
晴	炎热	高	强	取消	晴	寒冷	正常	弱	进行
阴	炎热	高	弱	进行	雨	适中	正常	弱	进行
雨	适中	高	弱	进行	晴	适中	正常	强	进行
雨	寒冷	正常	弱	进行	阴	适中	高	强	进行
雨	寒冷	正常	强	取消	阴	炎热	正常	弱	进行
阴	寒冷	正常	强	进行	雨	适中	正常	强	取消

参考答案：

- (1) 通常熵表示事物的混乱程度，熵越大表示混乱程度越大，越小表示混乱程度越小。
对于随机事件 S ，如果我们知道它有 N 种取值情况，每种情况发生的概论为，那么这件事的熵就定义为：

$$H(S) = - \sum_{i=1}^N P^{(i)} \log_2 P^{(i)}$$

信息熵：随机事件未按照某个属划的不同取值划分时的熵减去按照某个属性的不同取值

划分时的平均熵。即前后两次熵的差值。

(2)

$$H(\text{活动}) = -\left(\frac{9}{14} \log\left(\frac{9}{14}\right) + \frac{5}{14} \log\left(\frac{5}{14}\right)\right) = 0.94$$

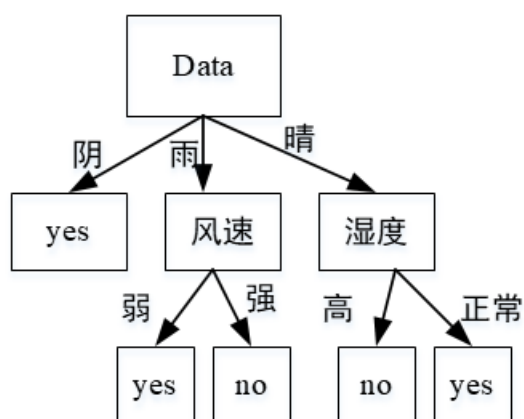
$$\text{Gain}(\text{活动}, \text{温度}) = 0.029$$

$$\text{Gain}(\text{活动}, \text{天气}) = 0.246$$

$$\text{Gain}(\text{活动}, \text{湿度}) = 0.151$$

$$\text{Gain}(\text{活动}, \text{风速}) = 0.048$$

因此选取天气属性作为划分属性；依次下去得到如下决策树



(3) 依据上面的决策图可判断可进行打垒球活动

五、(人工神经网络)(25 分)

构建一个有两个输入一个输出的单层感知器，实现对下表中的数据进行分类，设感知器的阈值为 0.6，初始权值均为 0.1，学习率为 0.6，误差值要求为 0，感知器的激活函数为硬限幅函数（阶跃函数），请给出迭代优化权值 w_1 和 w_2 的计算过程。（不超过 10 次迭代过程）（8 分）

X_1	X_2	d
0	0	0
0	1	0
1	0	0
1	1	1

(2) 请画出一个简单三层结构 BP 神经网络的结构图并写出正向计算和误差反向传播的公式，其输出层误差计算方式、两层权值修正方式均不限。（7 分）

(3) 请依据你对深度学习的了解，说明列举两种不同类型的深度神经网络模型，说明其基本原理、训练方法与应用领域？（10 分）

参考答案：(1)

对于第一个样本，输出神经元输入为

$$w_1(0)x_1(1) + w_2(0)x_2(1) = 0.1 * 0 + 0.1 * 0 = 0$$

输出神经元的输出为：

$$y(1) = f(0 - 0.6) = 0$$

权重调整：

$$w_1(1) = w_1(0) + \eta(d - y)x_1 = 0.1; \quad w_2(1) = 0.1$$

对于样本 2 和 3 同样；

对于样本 4，输出神经元输入为

$$w_1(3)x_1(4) + w_2(3)x_2(4) = 0.1 * 1 + 0.1 * 1 = 0.2$$

输出神经元的输出为：

$$y(4) = f(0.2 - 0.6) = 0$$

权重调整：

$$w_1(4) = w_1(3) + \eta(d - y)x_1 = 0.1 + 0.6(1 - 0) * 1 = 0.7; \quad w_2(4) = 0.7$$

到此完成一次循环，由于误差不为 0，则继续循环；

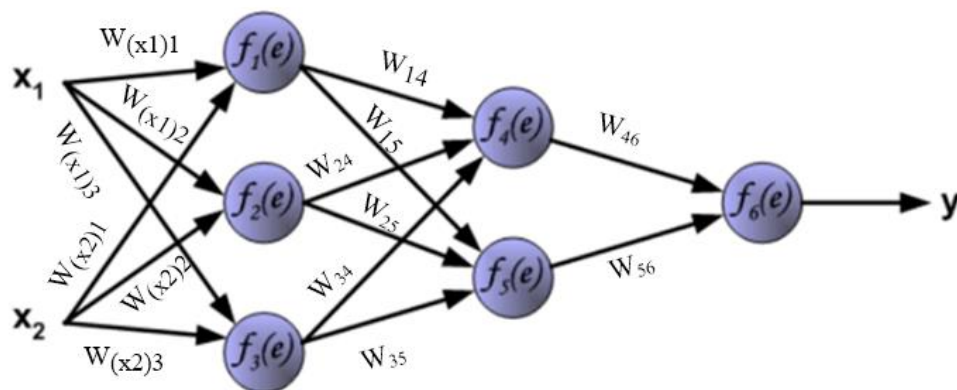
$$w_1(3)x_1(4) + w_2(3)x_2(4) = 0.7 * 1 + 0.7 * 1 = 1.4; \quad y(4) = f(1.4 - 0.6) = 1$$

此时误差为 0，终止循环。

从而

$$w_1(4) = w_1(3) + \eta(d - y)x_1 = 0.7 + 0.6(1 - 1) * 1 = 0.7 ; \quad w_2(4) = 0.7$$

(2)



第一层:

$$y_1 = f_1(w_{(x_1)1}x_1 + w_{(x_2)1}x_2); \quad y_2 = f_2(w_{(x_1)2}x_1 + w_{(x_2)2}x_2); \quad y_3 = f_3(w_{(x_1)3}x_1 + w_{(x_2)3}x_2)$$

第二层:

$$y_4 = f_4(w_{14}y_1 + w_{24}y_2 + w_{34}y_3); \quad y_5 = f_5(w_{15}y_1 + w_{25}y_2 + w_{35}y_3)$$

第三层:

$$y = f_6(w_{46}y_4 + w_{56}y_5)$$

误差传播:

输出层:

$$\sigma = z - y$$

第二层:

$$\sigma_4 = w_{46}\sigma; \quad \sigma_5 = w_{56}\sigma$$

第一层:

$$\sigma_1 = w_{14}\sigma_4 + w_{15}\sigma_5; \quad \sigma_2 = w_{24}\sigma_4 + w_{25}\sigma_5; \quad \sigma_3 = w_{34}\sigma_4 + w_{35}\sigma_5$$

以第一层为例，根据误差信号进行权值更新公式如下:

$$w'_{(x_1)1} = w_{(x_1)1} + \eta\sigma_1 \frac{df_1(e)}{de} x_1; \quad w'_{(x_2)1} = w_{(x_2)1} + \eta\sigma_1 \frac{df_1(e)}{de} x_2; \quad w'_{(x_1)2} = w_{(x_1)2} + \eta\sigma_2 \frac{df_2(e)}{de} x_1; \\ w'_{(x_2)2} = w_{(x_2)2} + \eta\sigma_2 \frac{df_2(e)}{de} x_2; \quad w'_{(x_1)3} = w_{(x_1)3} + \eta\sigma_3 \frac{df_3(e)}{de} x_1; \quad w'_{(x_2)3} = w_{(x_2)3} + \eta\sigma_3 \frac{df_3(e)}{de} x_2$$

(3) 卷积神经网络是人工神经网络的一种，已成为当前语音分析和图像识别领域的研究热点。它的权值共享网络结构使之更类似于生物神经网络，降低了网络模型的复杂度，减少了权值的数量。该优点在网络的输入是多维图像时表现的更为明显，使图像可以直接作为网络的输入，避免了传统识别算法中复杂的特征提取和数据重建过程。卷积网络是为识别二维形状而特殊设计的一个多层感知器，这种网络结构对平移、比例缩放、倾斜或者其他形式的变模型优缺点:

(1)优点:

(1)、权重共享策略减少了需要训练的参数，相同的权重可以让滤波器不受信号位置的影响来检测信号的特性，使得训练出来的模型的泛化能力更强；

(2)、池化运算可以降低网络的空间分辨率，从而消除信号的微小偏移和扭曲，从而对输入数据的平移不变性要求不高。

(2)缺点：

(1)、深度模型容易出现梯度消散问题。

在深度学习领域，传统的多层感知机为基础的上述各网络结构具有出色的表现，取得了许多成功，它曾在许多不同的任务上——包括手写数字识别和目标分类上创造了记录。但是，他们也存在一定的问题，上述模型都无法分析输入信息之间的整体逻辑序列。这些信息序列富含大量的内容，信息彼此间有着复杂的时间关联性，并且信息长度各种各样。这是以上模型所无法解决的，递归神经网络正是为了解决这种序列问题应运而生，其关键之处在于当前网络的隐藏状态会保留先前的输入信息，用来作当前网络的输出。

许多任务需要处理序列数据，比如 Image captioning, speech synthesis, and music generation 均需要模型生成序列数据，其他领域比如 time series prediction, video analysis, and musical information retrieval 等要求模型的输入为序列数据，其他任务比如机器翻译，人机对话，controlling a robot 的模型要求输入输出均为序列数据

模型优缺点：

(1)优点：

(1)、模型是时间维度上的深度模型，可以对序列内容建模；

(2)缺点：

(1)、需要训练的参数较多，容易出现梯度消散或梯度爆炸问题；

(2)、不具有特征学习能力。