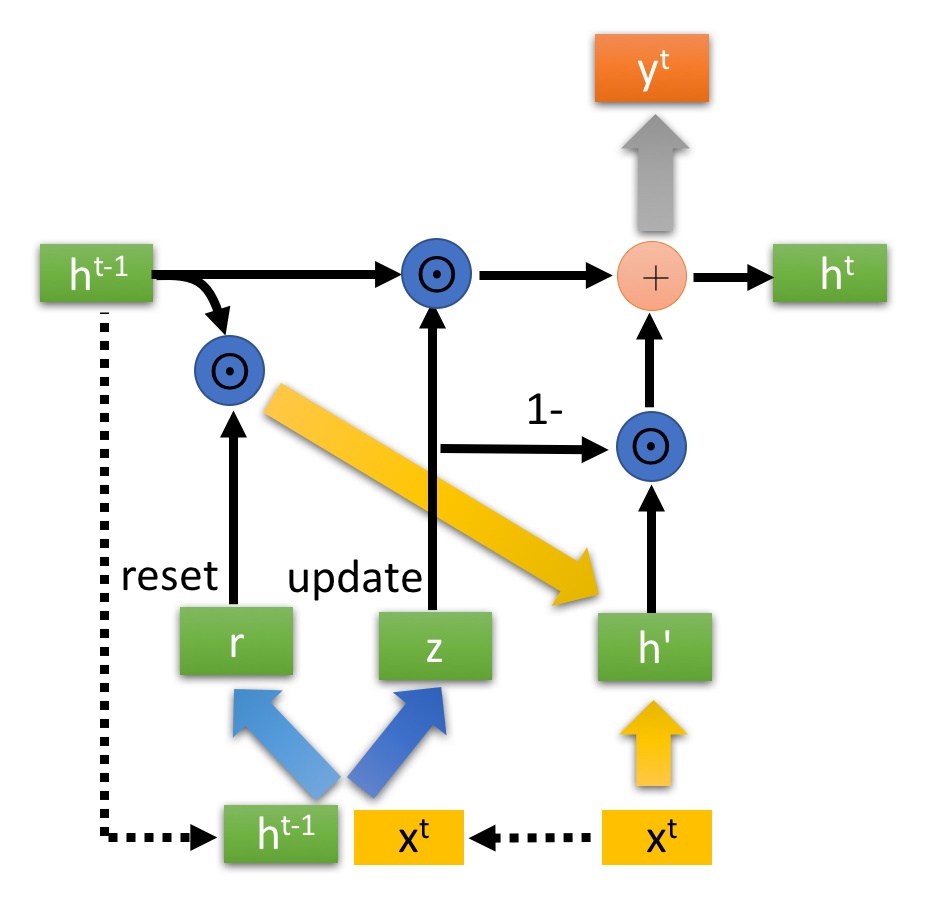
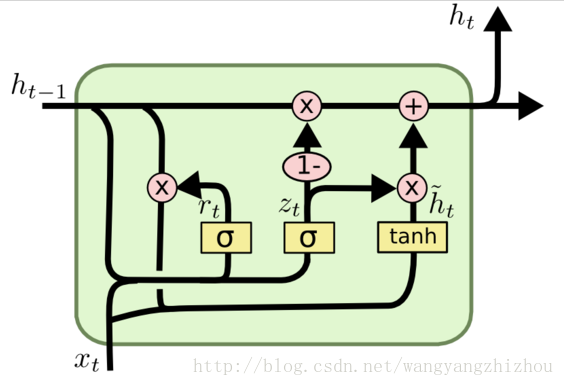
**GRU深入Gated Recurrent Unit**

**概述：**

GRU即Gated Recurrent Unit。前面说到为了克服RNN无法很好处理远距离依赖而提出了LSTM，而GRU则是LSTM的一个变体，当然LSTM还有有很多其他的变体。

GRU保持了LSTM的效果同时又使结构更加简单，易于计算，所以它也非常流行。

**内部结构：**

<https://zhuanlan.zhihu.com/p/32481747> 

【注】

\*\*\*\*计算基本对象：[ht-1,xt]拼接表示：过去+现在输入细胞状态表示:[batch,hid+input size]

它只有两个门了，分别为更新门和重置门，即图中的zt和rt。

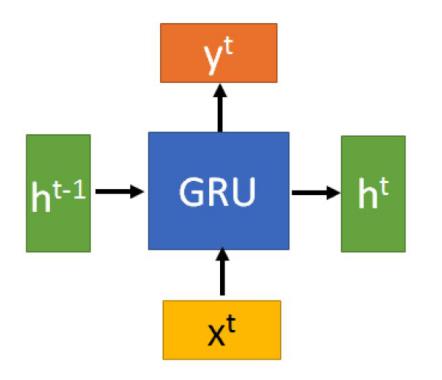
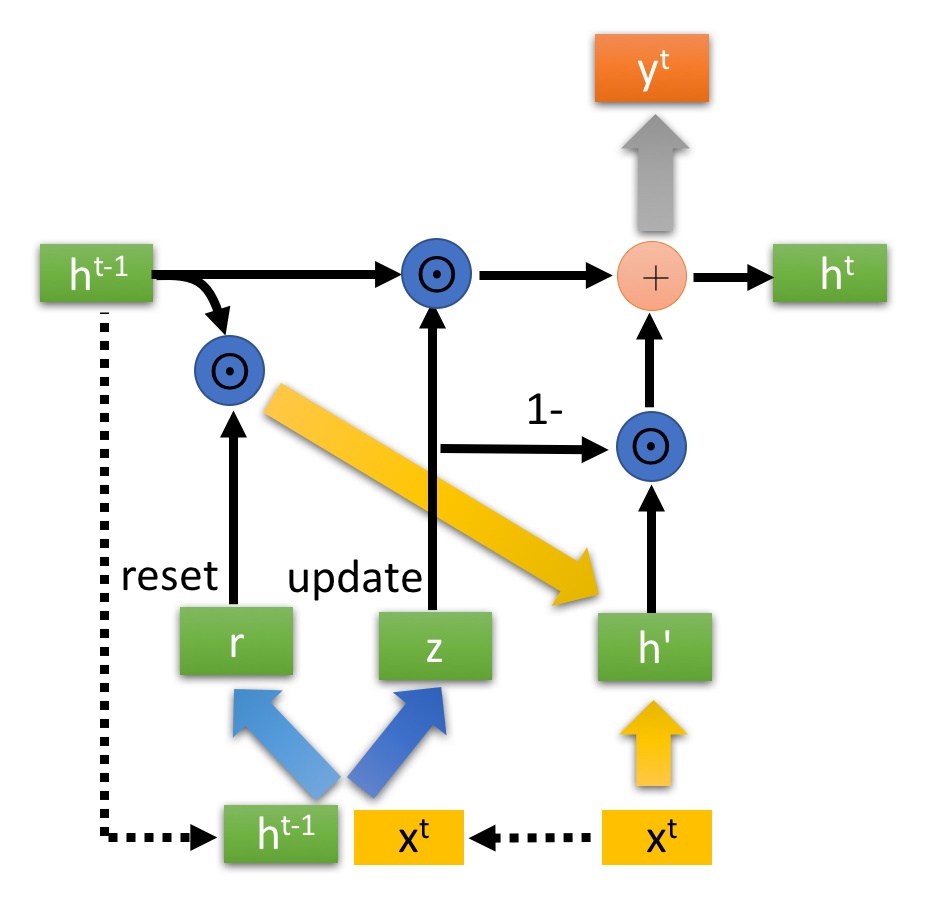
**（1）更新门（reset gate）Zt：**用于控制前一时刻的状态信息被带入到当前状态中的程度---**上一个状态的保留程度，**更新门的值越大说明前一时刻的状态信息带入越多。类似LSTM的输入门

----》上一个状态每个项的**保留程度**0-1

**（2）重置门（update gate）Rt：**用于控制忽略前一时刻的状态信息的程度---**上一个状态的遗忘程度**，重置门的值越小说明忽略得越多。类似LSTM中的遗忘门

----》上一个状态每个项的**遗忘程度**0-1

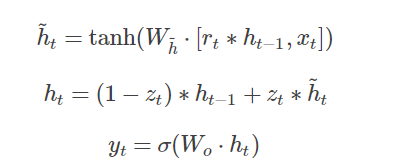
**输入输出结构：**

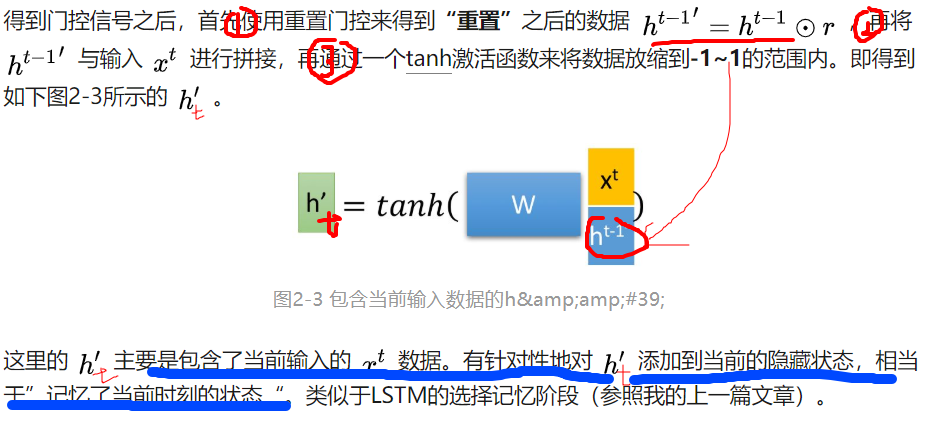
 

**输入：**Ht-1:上一个单元状态输出

Xt:当前单元输入 【batch,input size】

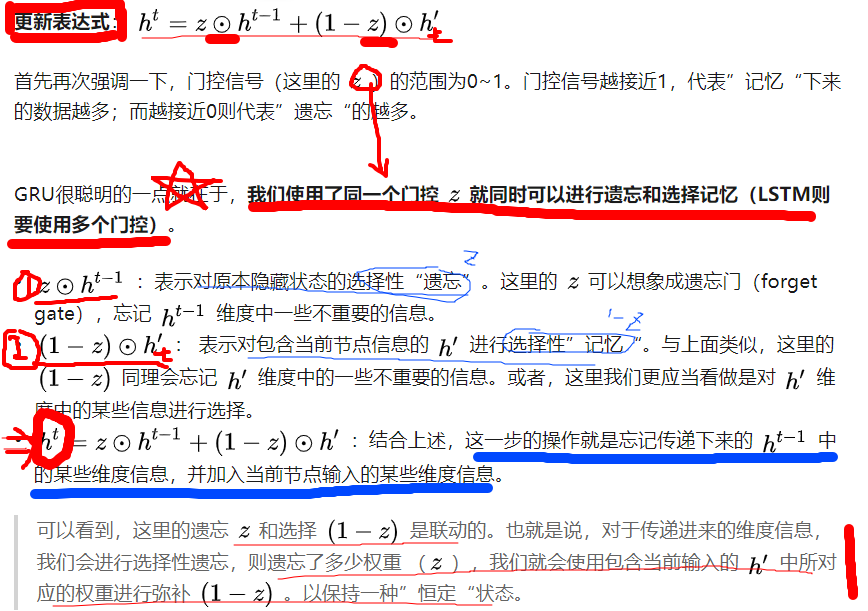
**输出：**





**关键：”更新记忆“阶段**

---》在这个阶段，我们**同时进行了遗忘和记忆两个步骤**。我们使用了先前得到的更新门控 IMG_256





**当前结点的输出：**

## **LSTM与GRU的关系**