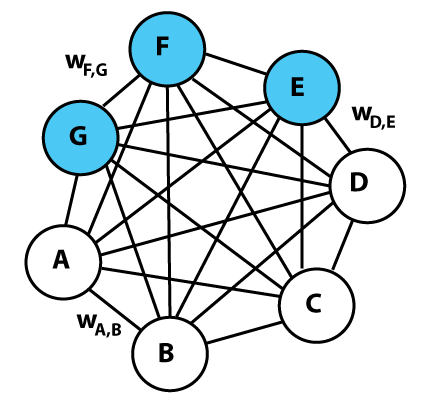
**Introdução**

Nas palavras de Bengio e Goodfellow, “Máquinas de Boltzmann Restritas são o exemplo quintessencial de como modelos de grafos são utilizados em *Deep Learning*”. Elas tiveram um papel importantíssimo no ressurgimento das redes neurais em 2006, permitindo o treinamento eficiente de modelos profundos. Dentre as aplicações mais notáveis de MBR estão redução de dimensionalidade, pré-treinamento de redes neurais, detecção de fraudes e sistemas de recomendação colaborativos. Elas inclusive foram usadas pelos campeões do [Prêmio Netflix](https://en.wikipedia.org/wiki/Netflix_Prize#Cancelled_sequel), numa competição em que a empresa pagaria 1 milhão de dólares para quem conseguisse melhorar o seu sistema de recomendação em 10%.

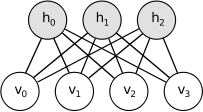
**Intuição**

Máquinas de Boltzmann são modelos probabilísticos (ou geradores) não supervisionados, baseados em energia. Isso significa que elas associam uma energia para cada configuração das variáveis que se quer modelar. Intuitivamente, aprender nesses modelos corresponde a associar configurações mais prováveis aos estados de menor energia. Nesses estados existem unidades que chamamos visíveis, denotadas por vv, e unidades ocultas, denotadas por hh. Para tornar isso mais correto, pense na Máquina de Boltzmann abaixo como representando os possíveis estados de uma festa. Cada ponto branco corresponde a uma pessoa que conhecemos e cada ponto azul, a uma que não conhecemos. Esses pontos assumem valor 1 se a pessoa for festa e 0 caso ela faltar. A probabilidade de uma pessoa ir à festa depende da probabilidade de todas as outras pessoas irem à festa.

Fonte: [Vera D, domínio público.](https://commons.wikimedia.org/w/index.php?curid=55007260)

Podemos pensar nas conexões como a relação entre as pessoas. Assim, a probabilidade de uma pessoa ir à festa depende dessas conexões, mas apenas se contabiliza as conexões das pessoas presentes (i.e. com 1 no ponto). Por exemplo, digamos que a conexão entre v1v1 e v4v4 seja negativa, indicando que essas pessoas não se gostam. Assim, se uma delas for a festa, a probabilidade da outra ir diminui. Note que algumas conexões podem ter valor próximo de zero, indicando que as pessoas são indiferentes entre si. Nesse caso, a presença de uma pessoa na festa não influencia diretamente na probabilidade de presença da outra, mas ainda pode haver influência indireta, por meio de outras pessoas. Por fim, há um estado da festa em que a maioria das pessoas presentes não se gostam. Esse estado é de bastante tensão ou **energia** e há uma tendência para que ele não ocorra frequentemente. As Máquinas de Boltzmann capturam isso ao colocar pouca probabilidade em estados com muita energia.

Com esse exemplo você já deve ter percebido que Máquinas de Boltzmann são extremamente complicadas. Afinal, para saber a probabilidade de que uma unidade esteja ligada (seja 1), é preciso saber o estado de outras, já que podem haver relações indiretas. De fato, Máquinas de Boltzmann são tão complicadas que ainda não mostraram utilidade prática. Por isso teremos que restringi-las de alguma forma. As Máquinas de Boltzmann Restritas cumprem esse papel. Elas são Máquinas de Boltzmann com a condição de que não há conexões diretas entre as unidades visíveis nem entre as ocultas. Isso as torna mais simples e práticas, mas também menos intuitiva; nosso exemplo da festa deixa de fazer muito sentido quando apenas pessoas conhecidas só interagem diretamente com pessoas desconhecidas. Em vez disso, infelizmente, terei que fornecer uma explicação intuitiva mais abstrata.

Fonte: [deeplearning.net](http://deeplearning.net/tutorial/rbm.html)

Apesar da restrição, Máquinas de Boltzmann Restritas, em teoria, conseguem representar qualquer fenômeno que quisermos, contanto que ela tenha unidades ocultas hh suficiente. As unidades visíveis nesse caso são as variáveis cuja interação entre si queremos entender. Com as MBR, forçamos a relação entre as unidades visíveis a acontecer indiretamente, por meio das unidades ocultas. Assim, quanto mais unidades ocultas, maior a capacidade da MBR de capturar interações complexas entre as variáveis.