**1.4 Машина Больцмана**

При решении технических и экономических задач неизвестна даже приблизительная оценка глобального экстремума. Это обуславливает применение методов глобальной оптимизации. НС Хопфилда находит локальный минимум задачи оптимизации. Для устранения этого недостатка можно использовать машину Больцмана, которая является расширением сети Хопфилда. В основе сети Больцмана лежит метод имитационного отжига (управляемого охлаждения), который является разновидностью процедуры случайного поиска. В своей базовой форме машина Больцмана является сетью Хопфилда и их структуры полностью совпадают. Метод имитации отжига представляет собой алгоритмический аналог физического процесса управляемого охлаждения. Он был предложен Метрополисом в 1953 году. Данный метод позволяет находить глобальный минимум функции нескольких переменных.

При отвердевании расплавленного металла его температура должна уменьшаться постепенно до момента полной кристаллизации. Если процесс остывания протекает слишком быстро, то образуются нерегулярности структуры металла, которые вызывают внутренние напряжения. В результате общее энергетическое состояние тела, зависящее от внутренней напряженности, остается более высоким, чем при медленном охлаждении. Быстрая фиксация энергетического состояния тела на уровне выше нормального соответствует сходимости оптимизационного алгоритма к точке локального минимума. Энергия состояния тела соответствует целевой функции, а абсолютный минимум – точке глобального минимума. Метод имитации отжига представляет собой алгоритмический аналог физического процесса управляемого охлаждения. Это метод позволяет находить глобальный минимум функции нескольких переменных.

Алгоритм обучения Больцмана:

* 1. Определить переменную T, представляющую искусственную температуру. Придать T большое начальное значение.
  2. Предъявить сети множество входов и вычислить выходы и целевую функцию.
  3. Придать случайное изменение  весу и пересчитать выход сети и изменение целевой функции в соответствии со сделанным изменением веса.
  4. Если целевая функция уменьшилась (улучшилась), то сохранить изменение веса. Если изменение веса приводит к увеличению целевой функции, то вероятность сохранения этого изменения вычисляется с помощью распределения Больцмана:

 (8)

где P(c) — вероятность изменения с в целевой функции; k — константа, аналогичная константе Больцмана, выбираемая в зависимости от задачи; T — искусственная температура.

Выбирается случайное число r из равномерного распределения от нуля до единицы. Если , то изменение сохраняется, в противном случае величина веса возвращается к предыдущему значению. Это позволяет системе делать случайный шаг в направлении, увеличивающем целевую функцию, и дает ей тем самым возможность вырываться из локальных минимумов, где любой малый шаг увеличивает целевую функцию.

Для завершения больцмановского обучения повторяют шаги 3 и 4 для **каждого** из весов сети, постепенно уменьшая температуру T, пока не будет достигнуто допустимо низкое значение целевой функции. В этот момент предъявляется другой входной вектор, и процесс обучения повторяется. Сеть обучается на всех векторах обучающего множества, с возможным повторением, пока целевая функция не станет допустимой для всех обучающих векторов.

Величина случайного изменения веса на шаге 3 может определяться различными способами. Например, подобно тепловой системе, весовое изменение может выбираться в соответствии с гауссовским распределением: где  — вероятность изменения веса на величину .

Машина Больцмана, учась на высокой температуре, ведет себя как случайная модель, а на низких температурах проявляет себя как детерминированная. Из-за случайной компоненты в процессе обучения, нейрон может принять новое значение состояния, которое увеличивается быстрее, чем уменьшается общее пространство состояний. Имитация физического отжига позволяет продвигаться к глобальному минимуму, избегая локальный. Для достижения сходимости к глобальному минимуму энергии скорость уменьшения температуры должна быть обратно пропорциональна логарифму времени.

Как и в сети Хопфилда, сети может быть представлен частичный образ для восстановления отсутствующей информации. Ограничение на число распознаваемых образов оценивается так же как и в сети Хопфилда классов – менее 15 % от общего количества элементов в слое.