объединенный
институт
ядерных
исследований
дубна

Р17-88-407

В.А.Загребнов, А.С.Чвыров

МОДЕЛЬ ЛИТТЛА - ХОПФИЛДА: РЕКУРРЕНТНЫЕ СООТНОШЕНИЯ ДЛЯ ВЕЛИЧИНЫ ОШИБКИ ВОССТАНОВЛЕНИЯ ОБРАЗОВ

Направлено в "Журнал экспериментальной
и теоретической физики"

1988

I°. Как теперь известно, сложная структура состояний спиновых стекол может быть использована для хранения и воспроизведения информации. В работах /I,2/ была предложена спиновая модель, которая работает как система ассоциативной памяти, т.е. является аналогом нейронных сетей мозга.

В этой модели каждый нейрон S_i, $i = 1,2,\dots,N$ находится в двух состояниях: $S_i = \pm 1$ (состояния возбуждения/торможения), индекс i пробегает сеть из N нейронов. Нейрон с индексом i находится под действием электрического поля U_i, потенциал которого определяется конфигурацией нейронов, связанных с i-тыми синаптическими связями V_{ij}:

$$U_i = \sum_{j\neq i}^{N} V_{ij}\, S_j\ . \tag{I}$$

Синаптические связи $\{V_{ij}\}$ могут быть возбуждающими $(V_{ij} > 0)$ или тормозящими $(V_{ij} < 0)$. Поэтому (в каждый данный момент) величины потенциалов $\{U_i\}_{i=1}^{N}$ являются функцией конфигурации нейронов $\underset{\sim}{S} = \{S_i\}_{i=1}^{N}$ и полного набора величины синаптических связей $\underset{\sim}{V} = \{V_{ij}\}$. Известно, что нейрон S_i переходит в состояние возбуждения /торможения, если потенциал превышает/ ниже некоторого порога $T_i^{\pm}$, т.е. условия локального равновесия имеют вид: $(U_i \mp T_i^{\pm}) S_i > 0$.

Модель Литтла - Хопфилда /2/ соответствует предположениям о том, что $T_i^{\pm} = 0$ (беспороговое возбуждение), $V_{ij} = V_{ji}$ (симметричные синапсы), а эволюция нейронной сети определяется релаксационной динамикой для нулевой температуры:

$$S_i(t+1) = \begin{cases} \operatorname{sign} U_i[\underset{\sim}{S}(t)] & ,\ U_i \neq 0 \\ S_i(t) & ,\ U_i = 0 \end{cases} \tag{2}$$

Тогда (2) соответствует процессу минимизации функции Ляпунова, которая в данном случае (см. (I)) имеет вид:

$$H[\underset{\sim}{S}] = -\sum_{1\le i<j\le N} V_{ij}\, S_i S_j\ . \tag{3}$$

А проблема запоминания образов (они кодируются "словами" $\{\underset{\sim}{\xi}^{(p)}\}_{p=1}^{M}$ в N бит и соответствуют неподвижным точкам $\{\underset{\sim}{S}_p^{*}\}_{p=1}^{M}$ динамики (2)) сводится к решению задачи о такой организации синаптических связей $\underset{\sim}{V}^{(M)}$, чтобы функция (3) (энергия, соответствующая заданной конфигурации нейронов $\underset{\sim}{S}$) имела не менее $2\cdot M$ глобальных минимумов.

Согласно гипотезе Хебба - Литтла - Купера /I-3/ образы $\{\underset{\sim}{\xi}^{(p)}\}_{p=1}^{M}$, поступающие в нейронную сеть в процессе обучения, модифицируют потенциалы синаптических связей по правилу:

$$V_{ij}^{(M)} = \frac{1}{N}\sum_{p=1}^{M} \xi_i^{(p)}\xi_j^{(p)}\ ,\quad i\neq j\ , \tag{4}$$

и являются случайными векторами с компонентами $\{\xi_i^{(p)} = \pm 1\}_{i=1}^{M}$. Такое правило обучения нейронной сети, вместе с релаксационной динамикой (2), действительно решает указанную выше проблему запоминания и воспроизведения нескоррелированных образов $\{\underset{\sim}{\xi}^{(p)}\}_{p=1}^{M}$: $N^{-1}\sum_{1\le i\le N} \xi_i^{(p)}\xi_i^{(p')} \sim O(N^{-1/2})$, когда их число M не очень велико: $\alpha = M/N \lesssim 0{,}14$, а температурные шумы в сети не превышают определенный уровень /4,5/.

При этих условиях нейронная сеть (3) функционирует как ассоциативная память. Процесс воспоминания образа $\underset{\sim}{\xi}^{(q)}$ сводится к эволюции начальной конфигурации $\underset{\sim}{S}(t=0)$ (см.(2)), релаксирующей к ближайшему (глобальному) минимуму гамильтониана, который соответствует этому образу. С ростом α интерференция образов, поступивших в нейронную сеть, приводит к тому, что "энергетический ландшафт", соответствующий (3), (4), становится сильно изрезанным. Для $\alpha = 1$ он соответствует модели Шерингтона - Киркпатрика /6/. В частности, глобальные минимумы, которые разделены барьерами (плато) с высотой $O(N)$ и соответствуют образам $\{\underset{\sim}{\xi}^{(p)}\}_{p=1}^{M}$, сохраняются только при небольших значениях α и имеют очень изрезанное дно /7/. Поэтому аттрактором для $\underset{\sim}{S}(t)$ будет не вектор $\underset{\sim}{\xi}^{(q)}$, соответствующий ближайшему к $\underset{\sim}{S}(t=0)$ глобальному минимуму, а некоторая его окрестность $A[\underset{\sim}{\xi}^{(q)}]$, определяемая структурой дна. Следовательно, перекрытие образа $\underset{\sim}{\xi}^{(q)}$ с воспоминанием $\underset{\sim}{S}(t)$:

$$m_N^{(q)}(t) = \frac{1}{N}\sum_{i=1}^{N} \xi_i^{(q)} S_i(t)\ , \tag{5}$$

при $t \to \infty$ не обязательно сходится к I (либо - I)[x)]. "Застревание" вектора $\underset{\sim}{S}(t)$ зависит также от деталей динамики нейронной сети.

Динамику (2) доопределяют обычно двумя способами /6,8/: последовательная (асинхронная) динамика, когда индекс i в (2) сканирует от I до N, и переворот очередного спина производится с учетом всех предыдущих переворотов; параллельная (синхронная) динамика, когда все спины переворачиваются/ непереворачиваются согласно (2) одновременно. Если $\underset{\sim}{S}(t=0)$ находится в области притяжения аттрактора $A[\underset{\sim}{\xi}^{(q)}]$, то первая динамика

x) Заметим, что в силу четности гамильтониана (3),(4) его глобальные минимумы по крайней мере дважды вырождены. Поэтому $\underset{\sim}{S}(t)$ может сходиться в окрестность "негатива" образа, т.е. вектора $-\underset{\sim}{\xi}^{(q)}$, если начальное перекрытие $m_N^{(q)}(t=0) < 0$.

(Монте - Карло для нулевой температуры T = 0) соответствует монотонной релаксации

$$H[\underset{\sim}{S}(t+1)] < H[\underset{\sim}{S}(t)] . \tag{6}$$

Локальные минимумы являются стационарными точками для асинхронной динамики. Поэтому воспоминание $\underset{\sim}{S}(t)$ может "застрять" достаточно далеко от $\underset{\sim}{\xi}^{(q)}$.

Синхронная динамика может нарушать (6) для отдельных моментов времени и обеспечивает релаксацию к $A[\underset{\sim}{\xi}^{(q)}]$ лишь на больших временных интервалах. Это делает "застревание" вектора $\underset{\sim}{S}(t)$ в локальных минимумах вдали от $A[\underset{\sim}{\xi}^{(q)}]$ менее вероятным.

Какая из динамик реализуется в нейронных сетях, неясно. В работах /2,4,6,8/ было уделено много внимания последовательной динамике. В недавних работах /6,8/ была сделана попытка проанализировать параллельную динамику для модели Литтла - Хопфилда. Однако полученные там результаты основаны на довольно грубых приближениях и находятся в сильном противоречии с численным экспериментом, который приведен в настоящей работе и работе /4/.

Целью настоящей работы являются : (а) вывод рекуррентных соотношений для эволюции перекрытия (5) в модели Литтла - Хопфилда, когда воспоминание $\underset{\sim}{S}(t)$ подчиняется паралельной динамике; (б) сравнение полученных результатов с численным экспериментом для параллельной динамики в этой модели.

2^0. Вывод рекуррентных соотношений для перекрытий начнем со следующего замечания. Содержательную теорию нейронных сетей удается построить только в пределе $N\to\infty$, $M\to\infty$, $M/N=\alpha$ /4,5,9/. Поэтому параллельную динамику (2) можно переопределить так, чтобы $S_i(t+1)=0$ для $U_i[\underset{\sim}{S}(t)]=0$, поскольку это не влияет на величины $m_\alpha^{(q)}(t)=\lim_{N\to\infty} m_N^{(q)}(t)$, которые и будут предметом нашего внимания ниже. Тогда с помощью (4) получаем x)

$$m_N^{(q)}(t+1)=\frac{1}{N}\sum_{i=1}^{N}\operatorname{sign}\left[m_N^{(q)}(t)+\xi_i^{(q)}U_i^{(q)}[\underset{\sim}{S}(t)]\right] . \tag{7}$$

x) Точнее, в (7) вместо $\{U_i^{(q)}[\underset{\sim}{S}(t)]\}_{i=1}^{N}$ должна стоять сумма (5) без i -го слагаемого, что вносит ошибку $O(N^{-1})$.

Здесь $\{U_i^{(q)}[\underset{\sim}{S}(t)]\}_{i=1}^{N}$ - последовательность (случайных) потенциалов, которые имеют вид:

$$U_i^{(q)}[\underset{\sim}{S}(t)] = \pm\frac{1}{N}\sum_{p\neq q}^{M}\sum_{j\neq i}^{N}\xi_i^{(p)}\xi_j^{(p)}\xi_j^{(q)}+\frac{1}{N}\sum_{p\neq q}^{M}\sum_{j\neq i}^{N}\xi_i^{(p)}\xi_j^{(p)}\left(S_j(t)\mp\xi_j^{(q)}\right) . \tag{8}$$

Здесь верхние знаки соответствуют случаю, когда $\underset{\sim}{S}(t=0)$ находится в области притяжения аттрактора $A[\underset{\sim}{\xi}^{(q)}]$, т.е. $m_N^{(q)}(t)>0$, а нижние знаки, когда начальное воспоминание находится в области притяжения "негатива", т.е. $A[-\underset{\sim}{\xi}^{(q)}]$, когда $m_N^{(q)}(t)<0$.

Рассмотрим подробнее первое слагаемое в (8). Для каждого $p=1,2,\ldots,M$ образ $\underset{\sim}{\xi}^{(p)}$ является реализацией последовательности независимых одинаково распределенных случайных величин $\underset{\sim}{\xi}(\omega)=\{\xi_i=\pm1\}_{i=1}^{N}$, $Pr\{\xi_i=\pm1\}=1/2$, т.е. $\underset{\sim}{\xi}^{(p)}=\underset{\sim}{\xi}(\omega_p)$. Поэтому в силу центральной предельной теоремы (см., например, /10/) суммы

$$\frac{1}{\sqrt{N}}\sum_{\substack{j=1\\(i\neq j)}}^{N}\xi_j^{(p)}\xi_j^{(q)}=\zeta_N^{(p)}$$

сходятся , при $N\to\infty$, к последовательности независимых (q -фиксировано и $p\neq q$) гауссовских случайных величин $\zeta^{(p)}$ с $M\zeta^{(p)}=0$ и $D\zeta^{(p)}=1$. По тем же причинам (с учетом независимости $\xi_i^{(p)}$ и $\zeta_N^{(p)}$) суммы

$$\frac{1}{\sqrt{M}}\sum_{\substack{p=1\\(p\neq q)}}^{M}\xi_i^{(p)}\zeta_N^{(p)}=\delta_i^{(M)} ,$$

при $M\to\infty$ сходятся к последовательности гауссовских случайных величин δ_i с $M\delta_i=0$ и $D\delta_i=1$. Поэтому в пределе $N\to\infty$, $M\to\infty$, $M/N=\alpha$ для первого слагаемого в (8) имеем:

$$\lim_{N\to\infty}\pm\frac{1}{N}\sum_{\substack{p=1\\(p\neq q)}}^{M}\xi_i^{(p)}\sum_{\substack{j=1\\(j\neq i)}}^{N}\xi_j^{(p)}\xi_j^{(q)}=\sqrt{\alpha}\,\delta_i^{\pm} , \tag{9}$$

где $\{\delta_i^{\pm}\}_{i\geq1}$ - последовательность одинаково распределенных независимых гауссовских случайных величин с $M\delta_i^{\pm}=0$ и $D\delta_i^{\pm}=\sigma_\delta^2=1$.

Аналогичные рассуждения справедливы и для второго слагаемого в (8). Необходимо лишь заметить, что последовательность $\eta_j^{\pm}=S_j\pm\xi_j^{(q)}$ принимает значения $\{0,\pm2\}$, причем $M\eta_j^{\pm}(t)=0$, $D\eta_j^{\pm}(t)=\lim_{N\to\infty}N^{-1}\sum_{1\leq j\leq N}(\eta_j^{\pm}(t))^2=2(1-|m_\alpha^{(q)}(t)|)$. Здесь для вычисления дисперсии мы воспользуемся эргодичностью последовательности $\{\eta_j^{\pm}(t)\}_{j\geq1}$. Далее, используя независимость $\underset{\sim}{\xi}^{(p)}$ и $\underset{\sim}{\eta}^{\pm}(t)$ и центральную предельную теорему, получаем:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{\substack{p=1\\(p\neq q)}}^{M}\xi_i^{(p)}\sum_{\substack{j=1\\(j\neq i)}}^{N}\xi_j^{(p)}\eta_j^{\pm}(t)=\sqrt{\alpha}\,\mu_i^{\pm}(t) . \tag{10}$$

Здесь $\{\mu_i^{\pm}(t)\}_{i\geq 1}$ – последовательность независимых одинаково распределенных гауссовских случайных величин с $M\mu_i^{\pm}(t)=0$ и $D\mu_i^{\pm}(t)=\sigma_\mu^2 = 2\,(1-|m_\alpha^{(q)}(t)|)$.

Таким образом, в указанном выше пределе последовательность $\{U_i^{(q)}[\underset{\sim}{S}(t)]\}_{i\geq 1}$ сходится к сумме гауссовских случайных последовательностей (9), (I0), коэффициент корреляции между которыми можно вычислить аналогично тому, как это было сделано для $D\eta_\delta^{\pm}$:

$$\tau = (\sigma_\delta\sigma_\mu)^{-1} M\,\delta_i^{\pm}\mu_i^{\pm}(t) =$$

$$= \lim_{N\to\infty}\left(N\sigma_\delta\sigma_\mu\right)^{-1}\sum_{i=1}^{N} \delta_i^{\pm}\mu_i^{\pm}(t) \simeq \left(\alpha\sigma_\delta\sigma_\mu\right)^{-1}\left(1-|m_\alpha^{(q)}(t)|\right)(1-\alpha) \; . \qquad \text{(II)}$$

Предположим, что последовательности $\{\xi_i^{(q)}\}_{i\geq 1}$ и $\{U_i^{(q)}[\underset{\sim}{S}(t)]\}_{i\geq 1}$ можно считать независимыми. Тогда в силу эргодической теоремы Биркгофа - Хинчина /I0/ и (9),(I0) рекуррентные соотношения (7), в пределе $N\to\infty$, $M/N=\alpha$, принимают вид:

$$m^{(q)}(t+1) = \int dx\, \frac{1}{2}\left\{\delta(x-1)+\delta(x+1)\right\} \iint du\, dv\; p_{\delta\mu}(u,v)\times$$

$$\times \operatorname{sign}\left[m^{(q)}(t)+x\sqrt{\alpha}\,(u+v)\right]. \qquad \text{(I2)}$$

Здесь $p_{\delta\mu}(u,v)$ – плотность вероятности распределения случайного потенциала $U_i^{(q)}$:

$$p_{\delta\mu}(u,v) = \left[2\pi\sigma_\delta\sigma_\mu\sqrt{1-\tau^2}\right]^{-1} \exp\left\{-\left[\frac{u^2}{\sigma_\delta^2} - 2\tau\frac{u\cdot v}{\sigma_\delta\sigma_\mu} + \frac{v^2}{\sigma_\mu^2}\right]\Big/ 2(1-\tau^2)\right\} . \qquad \text{(I3)}$$

Поэтому для рекуррентного соотношения, определяющего эволюцию перекрытия воспоминания и образа при паралелльной динамике, в пределе $N\to\infty$, $M/N=\alpha$, окончательно получаем:

$$m_\alpha^{(q)}(t+1) = \mathcal{F}_\alpha\left[m_\alpha^{(q)}(t)\right] \;, \quad q = 1, 2, \ldots \; . \qquad \text{(I4)}$$

Здесь

$$\mathcal{F}_\alpha[m] = \Phi\left(\frac{m}{\sqrt{\alpha+2(1-|m|)}}\right) \;,\quad \Phi(z) = \sqrt{\frac{2}{\pi}}\int_0^z dt\, \exp(-t^2/2) .$$

Нечетность функции $\mathcal{F}_\alpha[m]$ в правой части (I4) отражает упомянутое выше вырождение глобальных минимумов, соответствующее образам $\{\underset{\sim}{\xi}^{(q)}\}_{q\geq 1}$ и их "негативам" $\{-\underset{\sim}{\xi}^{(q)}\}_{q\geq 1}$. Поэтому ниже мы будем рассматривать только сходимость воспоминаний $\underset{\sim}{S}(t)$ к образам, т.е. $m_\alpha^{(q)}(t)\geq 0$.

3°. Из (I4) следует, что качество памяти существенно зависит от параметра насыщения нейронной сети α . При $\alpha < \alpha_c$ = 0,I398 рекуррентное соотношение (I4) имеет три неподвижные точки: две устойчивые $\bar{m}_\alpha = 0$, $m_\alpha^* > 0$ и одну $\tilde{m}_\alpha$ неустойчивую: $\bar{m}_\alpha < \tilde{m}_\alpha < m_\alpha^*$, см. рис. I и 2 . Точки $\bar{m}_\alpha$ и m_α^* являются аттракторами отображе-

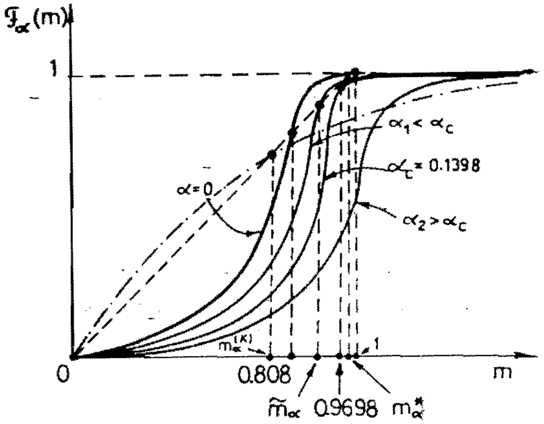

Рис. I. Графическое представление рекуррентного соотношения (I4) для различных значений параметра α . Точки $\bar{m}_\alpha = 0$ и m_α^* –аттракторы, точка $\tilde{m}_\alpha$ – сепаратриса. Штрих-пунктирная линия и $m_\alpha^{(K)}$ соответствуют рекуррентному соотношению (I6) для $\alpha < \alpha_c^{(K)}$.

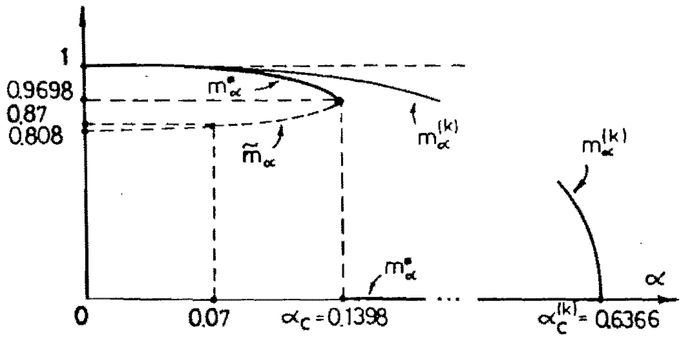

Рис. 2. Поведение неподвижных точек отображения (I4) как функций α : $\tilde{m}_{\alpha=0}$ = 0,808; $m^*_{\alpha_c}$ = 0,96978; α_c = 0,I398. То же для отображения (I6): $\alpha_c^{(K)}$ = 0,6366.

ния (I4), причем $m_\alpha^* \leq 1$ как раз и соответствует аттрактору $A[\underset{\sim}{\xi}^{(q)}]$, который вырождается в вектор $\underset{\sim}{\xi}^{(q)}$ только при $\alpha \to 0$, когда $m_\alpha^* \to 1$ x). Точка $\tilde{m}_\alpha$ является сепаратрисой отображения (I4). Наличие сепаратрис означает, что для сходимости воспоминания $S(t)$ в окрестность $A[\underset{\sim}{\xi}^{(q)}]$ образа $\underset{\sim}{\xi}^{(q)}$ необходимо, чтобы начальное перекрытие $m^{(q)}(t=0)$ превышало порог $\tilde{m}_\alpha > 0$. Заметим, что область притяжения аттрактора $A[\underset{\sim}{\xi}^{(q)}]$ довольно узка, причем порог не обращается в нуль при $\alpha \to 0$: $\tilde{m}_{\alpha\to 0} = 0{,}808$, см. рис. 2. (Предел $\alpha \to 0$ по-прежнему соответствует бесконечному числу образов, лишь $M/N \to 0$ при $N \to \infty$).

Если начальное воспоминание $\underset{\sim}{S}(t=0)$ таково, что $m^{(q)}(t=0) > \tilde{m}_\alpha$ и $m^{(p)}(t=0) \leq \tilde{m}$, для $p \neq q$, то

$$\lim_{t\to\infty} m_\alpha^{(p)}(t) = \begin{cases} m_\alpha^* & , p = q \\ 0 & , p \neq q \end{cases} \qquad (I5)$$

Для $\alpha > \alpha_c$ рекуррентные соотношения (I4) не имеют нетривиальных неподвижных точек: m_α^* скачком обращается в нуль, см. рис. 2. Это означает, что память переполнена образами и находится в состоянии "хаоса", т.е. не способна восстановить образ с ненулевым конечным перекрытием, даже если начальное воспоминание $\underset{\sim}{S}(t=0)$ близко к оригиналу.

Все эти результаты существенно отличаются от результатов работы Кинцеля [6], в которой для рекуррентного соотношения при <u>параллельной динамике</u> было получено выражение

$$m_\alpha^{(q)}(t+1) = \Phi\left(\frac{m_\alpha^{(q)}(t)}{\sqrt{\alpha}}\right) . \qquad (I6)$$

Из (I6) следует, что при $\alpha < \alpha_c^{(K)} = 2/\pi \simeq 0{,}6366$ имеются лишь две неподвижные точки: неустойчивая $\tilde{m}_\alpha = 0$ и устойчивая $m_\alpha^{(K)}$, для которой $m^{(K)}_{\alpha\to\alpha_c^{(K)}} \to 0$ <u>непрерывно</u> (см. рис. 2). При $\alpha \geq \alpha_c^{(K)}$, так же как и для (I4), имеется лишь одна неподвижная точка $\tilde{m}_\alpha = 0$.

В то же время, не зависящие от динамики термодинамические расчеты [4,II] для $T = 0$ дают $\alpha_c^{(A)} \simeq 0{,}138$ и скачок $m_\alpha^{(A)}$ в этой точке: $m^{(A)}_{\alpha_c^{(A)}-0} = 0{,}967$, $m^{(A)}_{\alpha_c^{(A)}+0} = 0$, ср. с m_α^* на рис. 2. Численный эксперимент, представленный в работе [4] (однако для <u>последовательной динамики</u>) подтверждает эти выводы. Начиная с $N = 3000$, отчетливо видно, что между $\alpha = 0{,}14$ и $\alpha = 0{,}16$ происходит скачкообразное уменьшение m_α с $m_{\alpha=0.14} \simeq 0{,}972$ до $m_{\alpha=0.16} \simeq 0{,}35$. Конечноразмерный скейлинг для $N = 500$, 1000, 2000, 3000 дает $\alpha_c^{(A)} \simeq 0{,}145$ [4].

В работе Кинцеля [6] также представлены результаты численных экспериментов. Они не согласуются с динамикой (I6). Например, они указывают на наличие порога и дают для его величины при $\alpha \simeq 0{,}075$ оценку $\tilde{m}_\alpha \simeq 0{,}4$, которая отличается от нашего результата $\tilde{m}_{\alpha=0} = 0{,}808$. Необходимо, однако, отметить, что в работе [6] также рассматривалась <u>последовательная динамика</u> xx) и, что более существенно, для небольшой системы: $N = 400$. Поэтому нами были проделаны численные эксперименты с параллельной динамикой, направленные прежде всего на проверку рекуррентного соотношения (I4) xx).

4°. Прежде, чем представить результаты численного эксперимента, сравним уравнение для неподвижных точек, которое следует из рекуррентного соотношения (I4):

$$m = \Phi\left(\frac{m}{\sqrt{\alpha + 2(1-|m|)}}\right) \qquad (I7)$$

с уравнением для параметра порядка перекрытия маттисовских состояний $(m^{(p)} = m\,\delta_{p,q})$:

$$m = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \xi_i^{(q)} \langle S_i \rangle$$

(здесь $\langle - \rangle$ - термодинамическое среднее с гамильтонианом (3) для нулевой температуры (см. [4,9,II]) :

$$m = \Phi\left(m/\sqrt{\alpha R}\right) ,$$
$$\sqrt{R} = 1 + \sqrt{2/\pi\alpha}\, \exp\left(-m^2/2\alpha R\right) . \qquad (I8)$$

---

x) Это означает, что $A[\underset{\sim}{\xi}^{(q)}] \to \underset{\sim}{\xi}^{(q)}$ при $\alpha \to 0$, только в том смысле, что $m_\alpha^* \to 1$: векторы $\{\hat{\underset{\sim}{\xi}}^{(s)}\}$, для которых $m^{(q)}(\hat{\underset{\sim}{\xi}}^{(s)}) = 1$, будем считать эквивалентными $\underset{\sim}{\xi}^{(q)}$.

x) В недавней работе [8] высказано мнение о том, что различие между динамиками не является существенным для свойств модели Литтла - Хопфилда.

xx) Специальный интерес мог бы представить численный эксперимент по сравнению параллельной динамики с релаксационной динамикой, которая рассматривалась в недавней работе [II].

Решения уравнения (I7) m_α и системы (I8) $m_\alpha^{(A)}$ очень близки: α_c = 0,I398, $\alpha_c^{(A)}$ = 0,I38, а $m^*_{\alpha_c}$ = 0,9698, $m^{(A)}_{\alpha_c^{(A)}}$ = 0,967. Поэтому, используя малость α ($<\alpha_c$) и близость нетривиального решения m_α к единице, из (I8) получаем:

$$m_\alpha = \Phi\left(m_\alpha \Big/ \sqrt{\alpha\left[1+2\sqrt{2/\pi\alpha}\,\exp(-m_\alpha^2/2\alpha)\right]}\right). \qquad (I9)$$

С другой стороны, в этой области изменения параметра α имеем:

$$1-|m_\alpha| = 1-\Phi(|m_\alpha|/\sqrt{\alpha}) \simeq \sqrt{2\alpha/\pi}\,\exp(-m_\alpha^2/2\alpha). \qquad (20)$$

Поэтому из (I9) и (20) получаем (I7), что и объясняет причину сходства решений m^*_α и $m_\alpha^{(A)}$, включая ветвь сепаратрис $\tilde{m}_\alpha$.

Численный эксперимент (на компьютере CDC - 6500, ОИЯИ -Дубна) был проведен для параллельной динамики и N = 6000. Такое большое N было выбрано для уточнения зависимости критических параметров от размеров системы. В работе /4/ отмечалось, что особенности поведения перекрытия m_α, характерные, например, для перехода памяти из маттисовских состояний, когда ошибка в воспроизведении образов мала, в фазу спинового стекла (хаотическое поведение памяти при переполнении образами) удается обнаружить только начиная с N = 2000.

В целом наш эксперимент подтвердил выводы работы /4/ (см. 3[0]), прояснил статус формул (I4) и (I7), а также некоторые особенности динамики: (а) Подтвердилось, что $0{,}I4 \leqslant \alpha_c < 0{,}I6$. Для α = 0,I4 воспоминание $\underset{\sim}{S}(t=0)$ с $m_\alpha(t=0)=1$ уже через I6 шагов сходилось к аттрактору, который в нашем случае состоял из двух векторов с $m_\alpha^{(1)}$ = = 0,9790 и $m_\alpha^{(2)}$ = 0,9797. Начиная с t = I6, вектор $\underset{\sim}{S}(t)$ становится периодической функцией: m_α (t = I6) = 0,9797, $m_\alpha(t=17)$= 0,9790, $m_\alpha(t=18)$= 0,9797 и т.д. При переходе от первого вектора аттрактора ко второму переворачивалось 8 спинов. Всего было сделано 26 итераций, см. рис. 3.

(б) Для α = 0,I6 перекрытие m_α, в отличие от формул (I7), (I8), не равно нулю. В нашем эксперименте m_α= 0,3457. Это значение достигалось после I98 итераций: m_α (t = I98÷206) = 0,3457. Для выяснения структуры аттрактора вычислялась удельная энергия $E(t)=H[\underset{\sim}{S}(t)]/N$. Оказалось, что аттрактор опять состоит из двух векторов: E (t =I98) = - 0,5843, E (t =I99)=- 0,58425, E (t = 200) = -0,5843 и т.д. Переход сопровождался переворотом I4 спинов. (в) Интересная особенность динамики была обнаружена после возмущения векторов аттрактора для α= 0,I4 с помощью шума (40 %). Это привело к уменьшению перекрытия зашумленного вектора $\underset{\sim}{S}_n$ (t = 0) с образом до m_α(t = 0) = = 0,I977. Эволюция $m_\alpha(t)$ для этого случая приведена на рис. 4. Она подтверждает наличие сепаратрисы (порога) с $\tilde{m}_{\alpha=0{,}14} \geqslant 0{,}4$, однако отли-

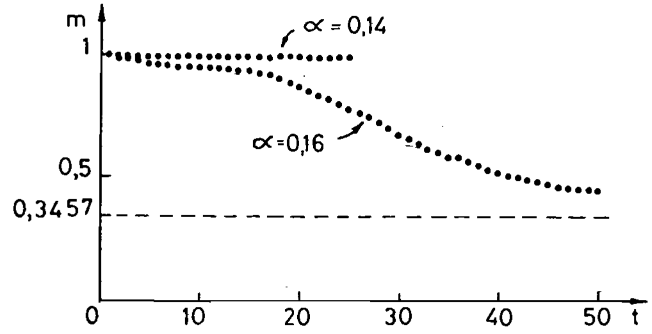

Рис. 3. Эволюция перекрытия $m_\alpha(t)$ (для α = 0,I4 и α = 0,I6) при начальном условии m_α(t = 0) = I.

чается от эволюции, предсказываемой формулой (I4). Главное отличие заключается в немонотонной зависимости $m_\alpha(t)$ и в том, что $\lim\limits_{t\to\infty} m_\alpha(t) > 0$. Вектор $\underset{\sim}{S}_n(t)$ вначале стремится к прообразу (идет процесс воспоминания), а лишь затем (после t = 2,3) начинает его забывать, но не окончательно, а с ненулевым предельным перекрытием m_α(t = 45) = 0,2090 и E (t = 45) = -0,5935. (г) Аналогичное поведение было обнаружено для эволюции при α = 0,I6 вектора $\underset{\sim}{S}_n(t)$, полученного из

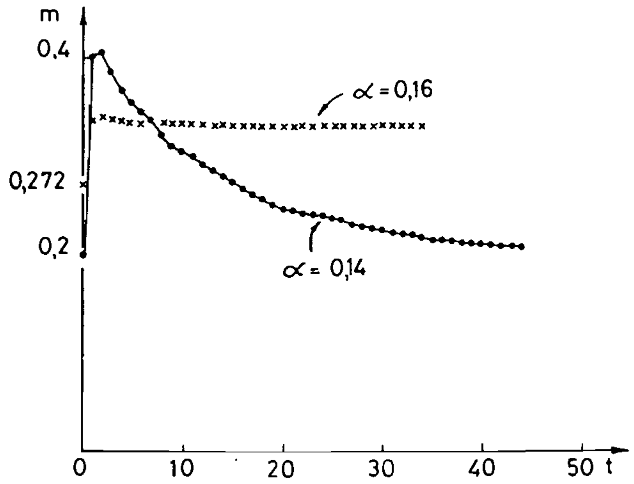

Рис. 4. Эволюция зашумлённых векторов для α= 0,I4 и α = 0,I6, когда начальное перекрытие $m_\alpha(t=0)$ ниже порога $\tilde{m}_\alpha$.

$\underset{\sim}{S}$ (t = 206) с помощью 10 % - го зашумления, которое привело к уменьшению перекрытия с образом до $m_\alpha(t=0)$ = 0,272, см. рис. 4. Было проделано 44 итерации, причем $m_\alpha(t = 44)$ = 0,3297, $E(t=44)$ = - 0,5865, ср. (б). (е) Для проверки гауссовости и независимости случайных величин, которые возникают при выводе (14) (см. п. 2°), для каждого момента t были построены соответствующие гистограммы. Если α = 0,14, то гистограммы для $\{U_i^{(q)}[\underset{\sim}{S}(t=0\div 25)]\}$ и $\{\xi_i^{(q)}U_i^{(q)}[\underset{\sim}{S}(t=0\div 25)]\}$ действительно близки к гауссовым с параметрами, близкими к предсказанным в п. 2°, см. рис. 5 . Для α = 0,16 они лишь для нескольких первых шагов близки к гауссовским, а затем существенно от них отличаются, см. рис.6.

5°. <u>Выводы</u>. Приближенное рекуррентное соотношение (14) хорошо описывает эволюцию модели Литтла - Хопфилда только для $\alpha \leqslant \alpha_c$, причем только в окрестности аттракторов. Отклонение отображения (14) от численного эксперимента (особенно для $\alpha > \alpha_c$, см. провал на гистограмме рис. 6.) - следствие пренебрежения корреляционными эффектами и прежде всего между $\{\xi_i^{(q)}\}$ и $\{U_i^{(q)}\}$, см. 2° и обсуждение в работах /6,8/ . При этом формула (17) для неподвижных точек с большой точностью воспроизводит решения системы (18), полученной из термодинамических соображений для T = 0. Значение α_c = 0,1398, полученное из (17), ближе к тому, что дает численный эксперимент: $\alpha_c \simeq$ 0,145, чем $\alpha_c^{(A)}$ = 0,138, которое следует из уравнений (18). Существенно,

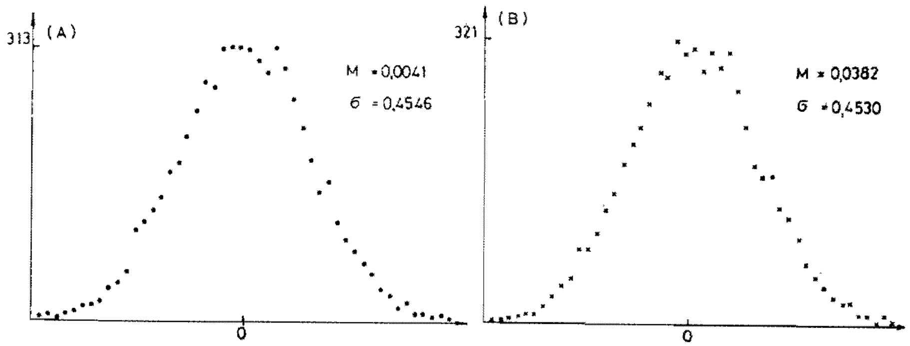

Рис. 5. (А) Гистограмма $\{U_i^{(q)}[\underset{\sim}{S}(t=25)]\}$ для α = 0,14
(В) Гистограмма $\{\xi_i^{(q)}U_i^{(q)}[\underset{\sim}{S}(t=25)]\}$ для α = 0,14
Средние и дисперсии указаны на рисунке.

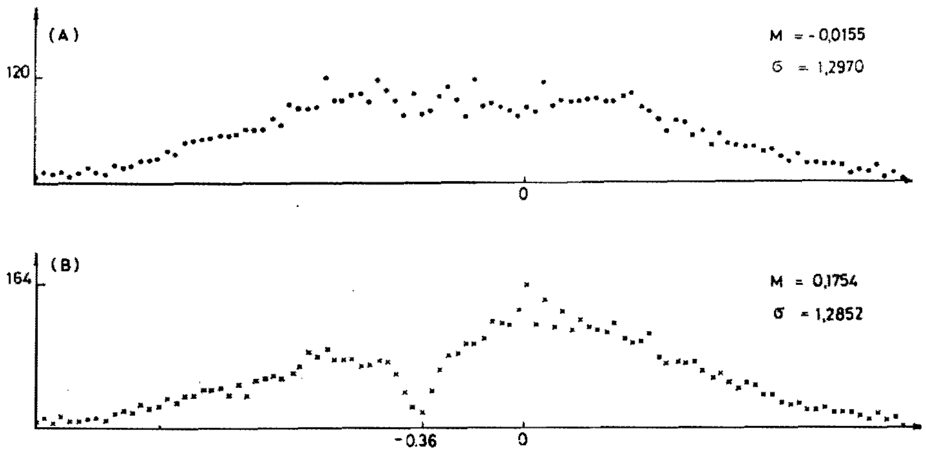

Рис. 6. (А) Гистограмма $\{U_i^{(q)}[\underset{\sim}{S}(t=206)]\}$ для α = 0,16
(В) Гистограмма $\{\xi_i^{(q)}U_i^{(q)}[\underset{\sim}{S}(t=206)]\}$ для α = 0,16
Средние и дисперсии указаны на рисунке.

что рекуррентное соотношение (14) настоящей работы предсказывает наличие порога $\tilde{m}_\alpha > 0$, начиная с которого,происходит "захват" воспоминания $\underset{\sim}{S}(t)$ аттрактором соответствующего образа. Величина $\tilde{m}_\alpha$=0,808,полученная из (14),(17), по-видимому,завышена.Все эти результаты отличаются от тех,которые дает динамика,предложенная Кинцелем (16): отсутствие скачка параметра m_α^* , отсутствие порога: $\tilde{m}_\alpha$ = 0 и т.д. Напомним, что численный эксперимент, проведенный в работе /6/, для N=400 и <u>последовательной динамики</u> дает $\tilde{m}_{\alpha=0,075} \simeq$ 0,4.

Наконец, для α = 0,16 > α_c обе формулы (17) и (18) дают m_α^* = 0, что не соответствует численному эксперименту. Нулевое значение $m_{\alpha>\alpha_c}$ отвечает тому, что при переполнении памяти образами (при $\alpha > \alpha_c$) устойчивым является <u>состояние спинового стекла</u>, а не <u>маттисовские состояния.</u> Обсуждение физического смысла этого остаточного перекрытия можно найти в работах /4-7/. Наличие остаточного перекрытия для $\alpha < \alpha_c$ при эволюции вектора $\underset{\sim}{S}(t)$ с начальным условием ниже порога $\tilde{m}_\alpha$ впервые обнаружено в настоящей работе. Оно соответствует "застреванию" вектора $\underset{\sim}{S}(t)$ в локальном минимуме на изрезанном плато "энергетического ландшафта" вдали от аттракторов, которые отвечают образам, хранящимся в памяти. Последние представляют собой систему глобальных минимумов $\{A[\underset{\sim}{\xi}^{(q)}]\}$, разделенных плато, см. п. 1° и 2°.

ЛИТЕРАТУРА

I. Little W.A. Math. Biosci. 1974, 19, 101; 1978, 39,281.
2. Hopfield J.J. Proc. Nat. Acad. Sci USA, 1982, 79, 2554; 1984, 81, 3088.
3. Hebb D.O. The Organization of Behavior. Wiley. New York, 1949.
4. Amit D.J., Gutfreund H., Sompolinsky H. Phys. Rev. Lett., 1985, 55, 1530; Ann. of Phys., 1987, 173, 30.
5. Ioffe L.B. Feigel'man M.V. Evrophys. Lett., 1986, 1, 197.
6. Kinzel W.Z. für Phys., 1985, B60, 205.
7. Newman Ch. M. Memory Capacity in Neural Networks Model : Rigorous Lower Bounds, Preprint, 1987.
8. Bruce A.D., Cardner E.J., Wallace D.J. of Phys. A., 1987, A20, 2909.
9. Van Hemmen J.L., Zagrebnov V.A. J. of Phys. A., 1987, A20, 3899.
10. Гнеденко Б.В.Курс теории вероятностей. М.: Наука, 1988.
11. Feigel'man M.V. Ioffe L.B., Int. J. of Modern Phys. B., 1987, 1, 51.

Рукопись поступила в издательский отдел
7 июня 1988 года.

Загребнов В.А., Чвыров А.С. Р17-88-407

Модель Литтла – Хопфилда: рекуррентные соотношения для величины ошибки восстановления образов

Для случая параллельной /синхронной/ динамики выведены приближенные рекуррентные соотношения для величины перекрытия воспоминаний и истинных образов. Критическое значение параметра насыщения нейронной сети $a = M/N/M$ – число образов, а N – число нейронов/, полученное из этих соотношений, равно $a_c = 0,1398$ в пределе $N \to \infty$. Проведено сравнение с соответствующими термодинамическими расчетами и численным экспериментом для N = 6000.

Работа выполнена в Лаборатории теоретической физики и в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна 1988

Перевод О.С.Виноградовой

Zagrebnov V.A., Chvyrov A.S. P17-88-407

Little-Hopfield Model: Recursion Relations for Retrieval Pattern Errors

For zero temperature parallel dynamics approximate recursion relations for the pattern-retrieval-pattern-overlap parameter is derived. They give for the critical value of the neutral network saturation parameter: a_c = =0.1398. The result is compared with thermodynamic calculations for T = 0 and numerical simulations for a network with the size N = 6000.

The investigation has been performed at the Laboratory of Theoretical Physics and at the Laboratory of Computing Techniques and Automation, JINR.

Preprint of the Joint Institute for Nuclear Research. Dubna 1988