

Теория адаптивного резонанса Гроссберга

Александр Панов

ИСА РАН

13 февраля 2015 г.

- Стефан Гроссберг, 75 лет — специалист по когнитивным наукам, нейрофизиолог, математик.
- Профессор Бостонского университета.
- Scopus: более 380 статей, h-индекс — 64.
- Работы с количеством цитирований более 1000:

Основные проблемы

- Проблема соотношения стабильности и пластичности в восприятии.
- Принцип дополняющих когнитивных процессов.
- Принцип послышной организации всех когнитивных процессов.
- Все состояния сознания являются состояниями резонанса.

WHAT

Spatially-invariant object learning and recognition

Fast learning without catastrophic forgetting

IT

WHERE

Spatially-variant reaching and movement

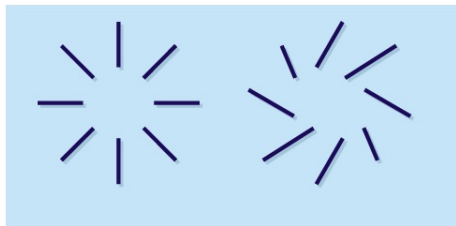
Continually update sensory-motor maps and gains

PPC

MATCHING

LEARNING

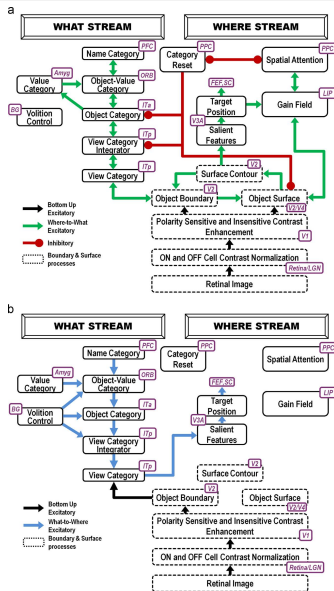
WHAT	WHERE
EXCITATORY	INHIBITORY
MATCH	MISMATCH



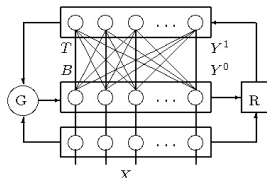
Что, где и как в коре головного мозга

Дорсальный (Что?) и вентральный (Где?, Как?) пути обработки информации.

- Retina/LGN — сетчатка/боковое коленчатое ядро.
- V1, V2, V3, V4 — первичная и вторичные зоны зрительной коры.
- IT — нижневисочная (инферотемпоральная) зона коры.
- ORB — глазнично—лобная (орбитофронтальная) зона коры.
- LIP — латеральная межтеменная (интрапариетальная) зона коры.
- PPC — задняя теменная (париетальная) зона коры.
- PFC — передняя лобная зона коры.
- FEF — глазодвигательная кора.

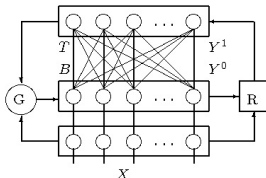


Нейронная сеть ART-1



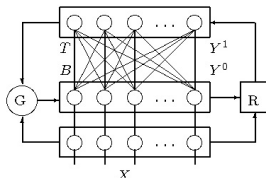
- Четыре функциональных модуля: слой сравнения, слой распознавания и два управляющих нейрона: сброса R и управления G .
- Вектора X , Y^0 , Y^1 — двоичные вектора.
- Для нейронов слоя сравнения используется правило 2/3.

Нейронная сеть ART-1



- **Начальная фаза:** $G = 0$ при $X = \{0, \dots, 0\}$; при поступлении ненулевого X принимаем $G = 1$, $Y^0 = X$ и определяем $Y^1 = \{y_1^1, \dots, y_n^1\}$, $\forall i \neq w \ y_i^1 = 0, y_w^1 = 1$, где w — индекс нейрона-победителя ($\max_j \sum_i y_i^0 * b_{ij}$) в слое распознавания, обнуляем управление $G = 0$.
- **Фаза сравнения:** Новый выход слоя сравнения $Y^0 = \{t_{1w}, \dots, t_{mw}\} \wedge X$; если $\|Y^0\|/\|X\| > \rho$, то завершаем процесс, иначе возникает сигнал сброса, подавляется нейрон-победитель и начинается фаза поиска.
- **Фаза поиска:** так как нейрон победитель подавлен, то R обнуляется, $G = 0$, $Y^0 = X$, определяется новый нейрон-победитель и снова переходим к фазе сравнения.

Нейронная сеть ART-1

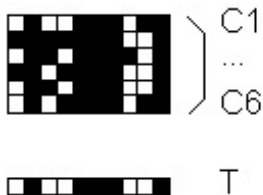


Окончание итерационного процесса:

- Найдётся запомненная категория, сходство которой с входным вектором X будет достаточным для успешной классификации. После этого запускается **фаза обучения**, в котором модифицируются веса b_{iw} и t_{iw} матриц B и T для победившего нейрона.
- В процессе поиска все запомненные категории окажутся проверенными, но ни одна из них не дала требуемого сходства. В этом случае входной образ X объявляется новым для нейросети, и ему выделяется новый нейрон в слое распознавания.

Фаза обучения

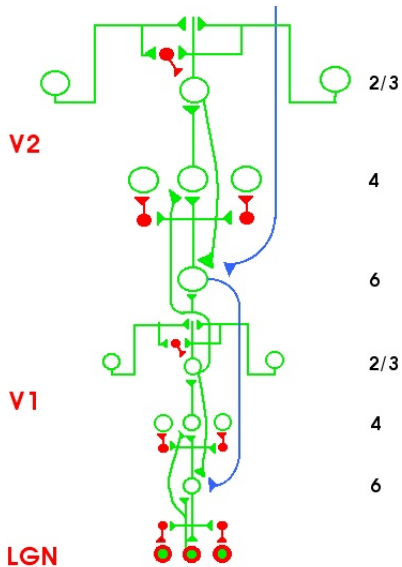
- Начальные значения матриц B и T должны удовлетворять соотношениям: $b_{ij} < L/(L - 1 + m)$, $t_{ij} = 1$, m — размерность входных данных (вектора X).
- При нахождении подходящей категории по вектору $Y^0 = \{y_1^0, \dots, y_m^0\}$ $b_{ij} = (L * y_i^0)/(L - 1 + \sum_k y_k^0)$, $\forall i$ $t_{ij} = y_i^0$.
- Обучение, таким образом, сопровождается занулением все большего числа компонент матрицы T , оставшиеся ненулевыми компоненты определяют множество критических черт найденных категории.



Некоторые свойства модели

- По достижении стабильного состояния в результате обучения предъявление входного вектора будет сразу приводить к правильной классификации без фазы поиска, на основе прямого доступа.
- Процесс поиска устойчив.
- Процесс обучения устойчив. Обучение весов нейрона—победителя не приведёт в дальнейшем к переключению на другой нейрон.
- Процесс обучения конечен. Итоговое состояние для заданного набора образов будет достигнуто за конечное число итерации, при этом дальнейшее предъявление этих образов не вызовет циклических изменений значений весов.

Принцип многослойности



Дальнейшее развитие

- ART-2 — переход от двоичных значений к непрерывным (действительным).
- ART-3 — добавление нескольких слоёв для сжатия информации и абстрагирования. Резонанс возникает на некотором уровне иерархии. Введён механизм рефрактерного торможения.
- FuzzyART — введение элементов нечёткой логики.
- ARTMAP, рART — комбинация двух блоков ART-1 и ART-2 для регулировки параметра сброса (чувствительности).
- Много прикладных моделей: dARTEX, ARTPHONE и др.