

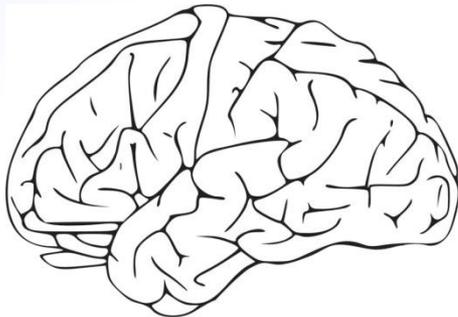
Интеллектуальная система поддержки принятия решений — применяемые алгоритмы и способ реализации

Докладчик:

Бухаров О.Е.

Аспирант 3 г.о.

МИЭМ НИУ ВШЭ



Научный руководитель:
доцент Боголюбов Д.П.

Прогнозирование площади морских льдов в Северном Полушарии

- Среднесрочное прогнозирование
- Проверка гипотезы о влиянии приходящей солнечной радиации на площадь льдов



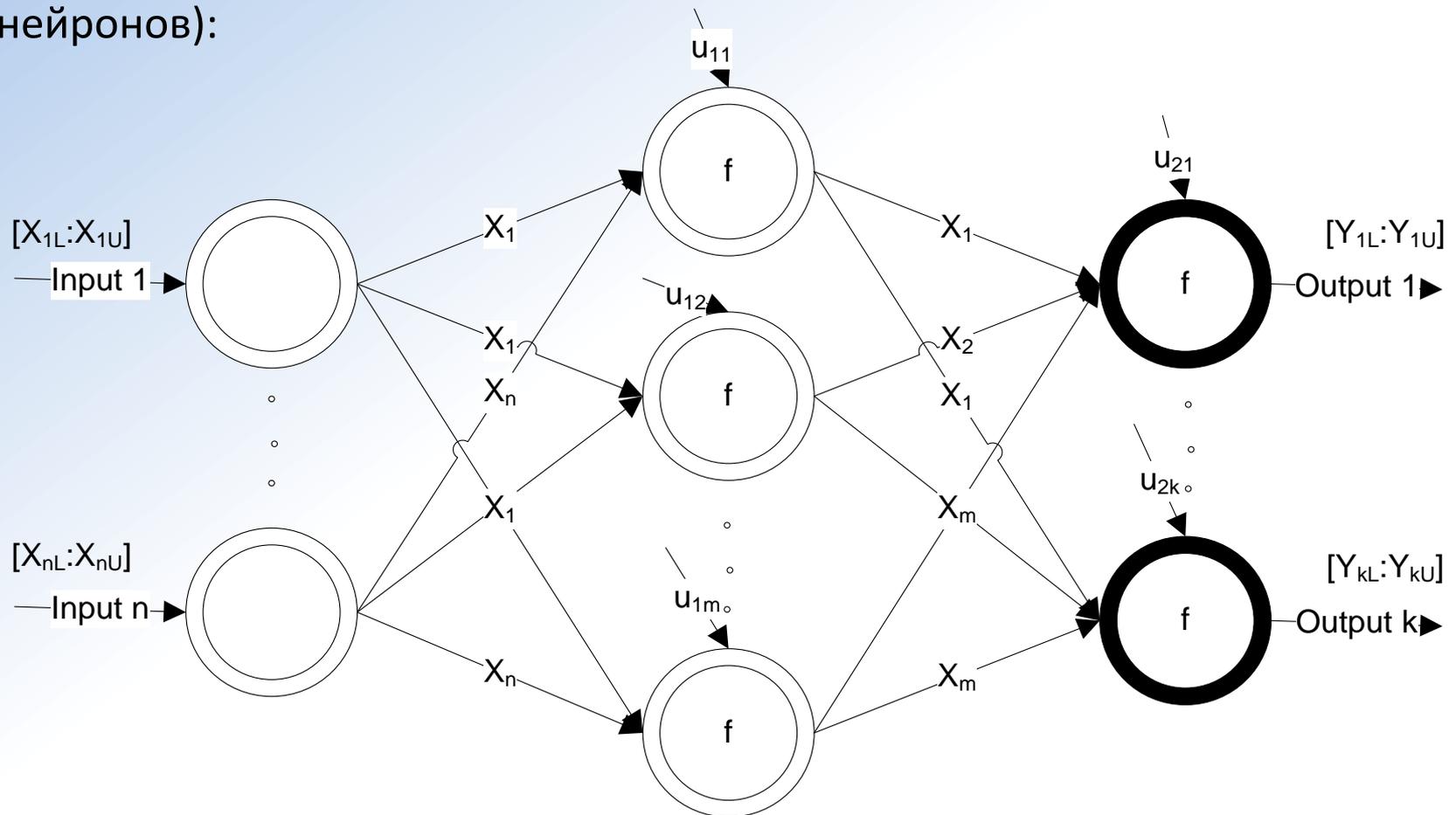
Цель работы

Разработка ИСППР, удовлетворяющей следующим требованиям:

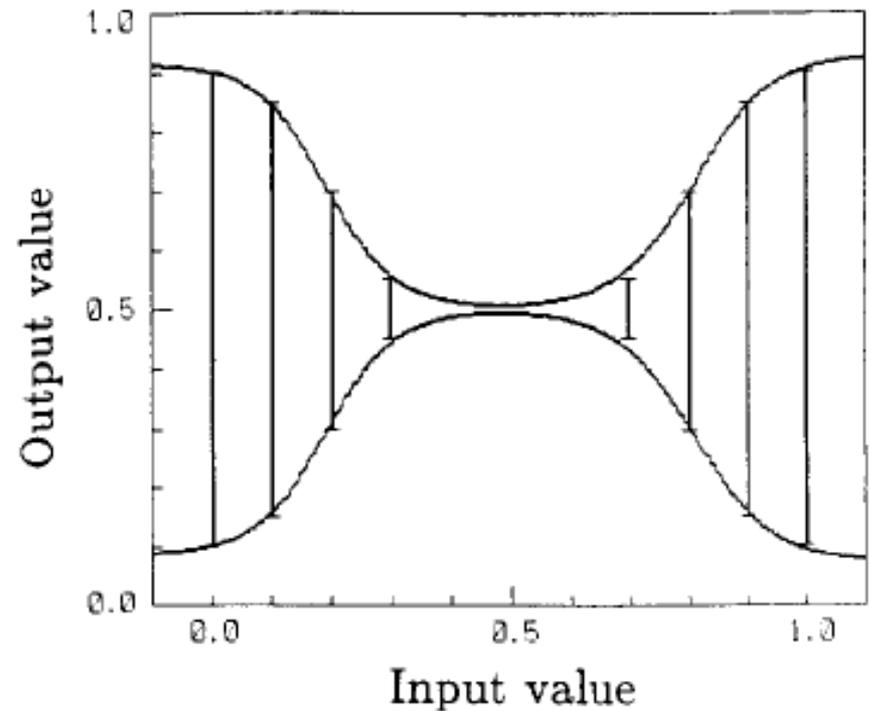
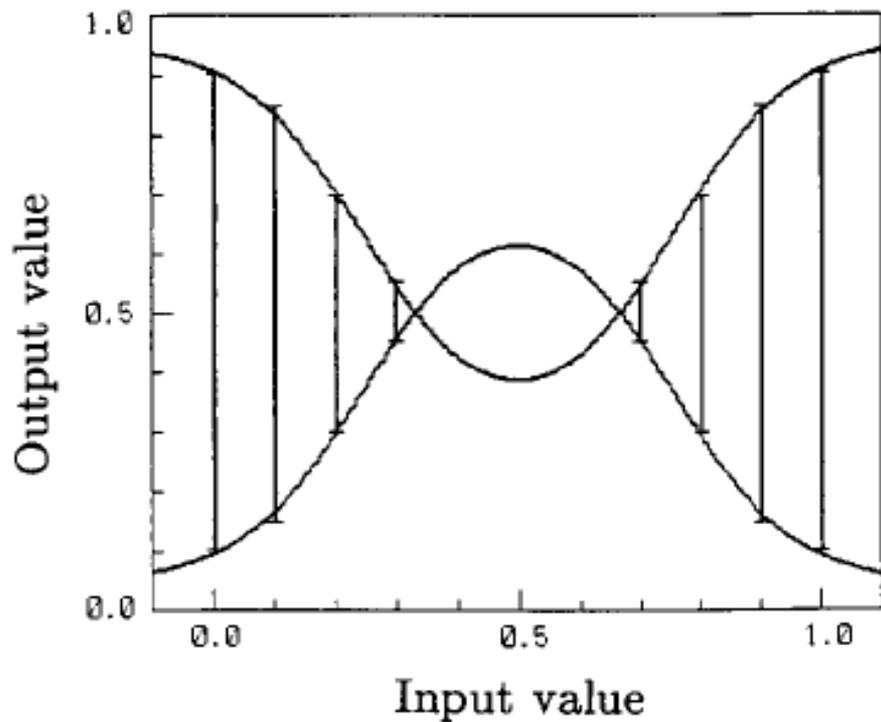
- оперативная оценка зависимости между входными параметрами и прогнозируемой величиной в сложных системах;
- уменьшение требований к квалификации пользователя,
- возможность работы с недоопределенными данными,
- отсутствие привязки к конкретной предметной области.

Интервальная нейронная сеть

Схема интервальной нейронной сети с одним скрытым слоем (n – число входных нейронов, m – число скрытых нейронов, k – число выходных нейронов):



Прогнозирование интервалов двумя стандартными сетями



Справа интервальная нейронная сеть, слева две стандартные нейронные сети. Вертикальными линиями изображены выходные интервалы обучающих наборов.

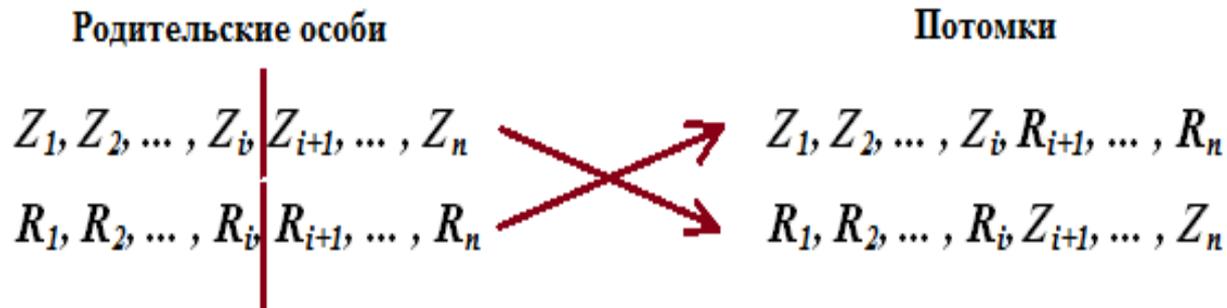
Преимущества нейронных сетей

- Нейронная сеть – сложная многомерная нелинейная регрессия
- Возможность работы с неинформативными, шумовыми входными сигналами.
- Нейронная сеть имеет меньшие требования к квалификации человека, использующего ее, чем сложные статистические модели способные давать аналогичные результаты.
- Нейронная сеть может одновременно решать несколько задач при наличии у неё нескольких выходов.
- Существуют алгоритмы, позволяющие решать обратные задачи на нейронной сети, обученной для решения прямой задачи.
- Изначально задав синаптические веса нейронной сети можно воссоздать и проверить предполагаемые статистические модели, а также улучшить их путем тренировки сети.

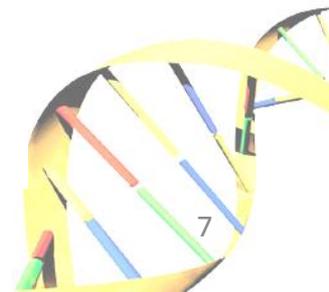
Генетический алгоритм

Генерация случайного начального поколения

1. Отбор наилучших особей для создания нового поколения
2. Скрещивание отобранных особей
3. Оценка нового поколения

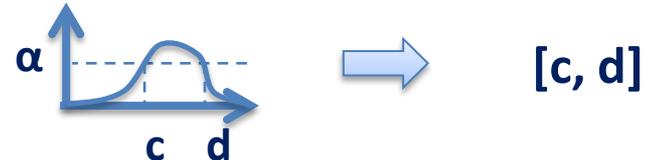
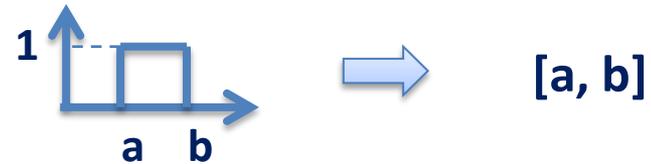


Процедура скрещивания



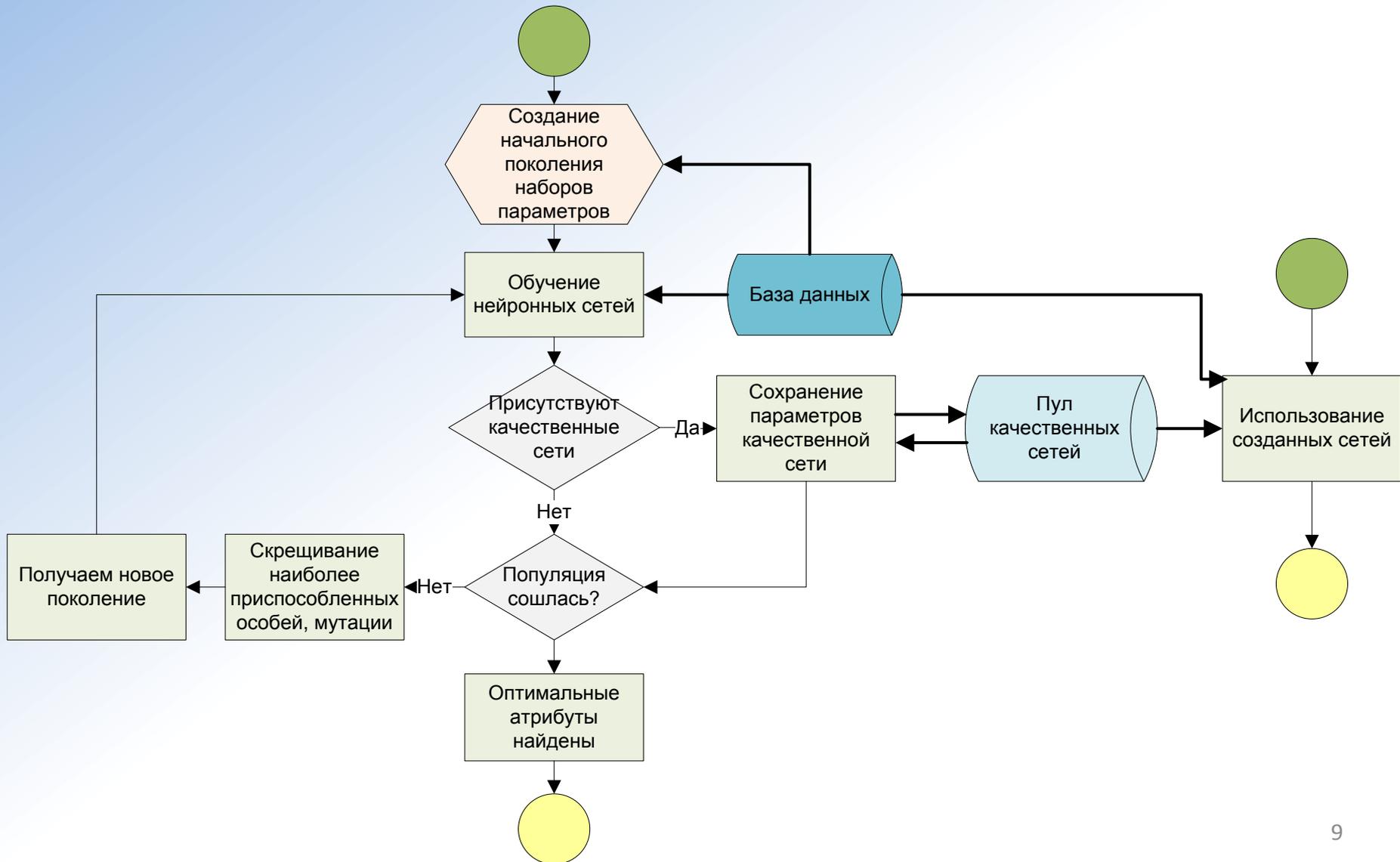
Модуль работы с нечеткими переменными

- Определение возможных значений для нечеткой переменной
- Определение простых функций принадлежности для каждого значения

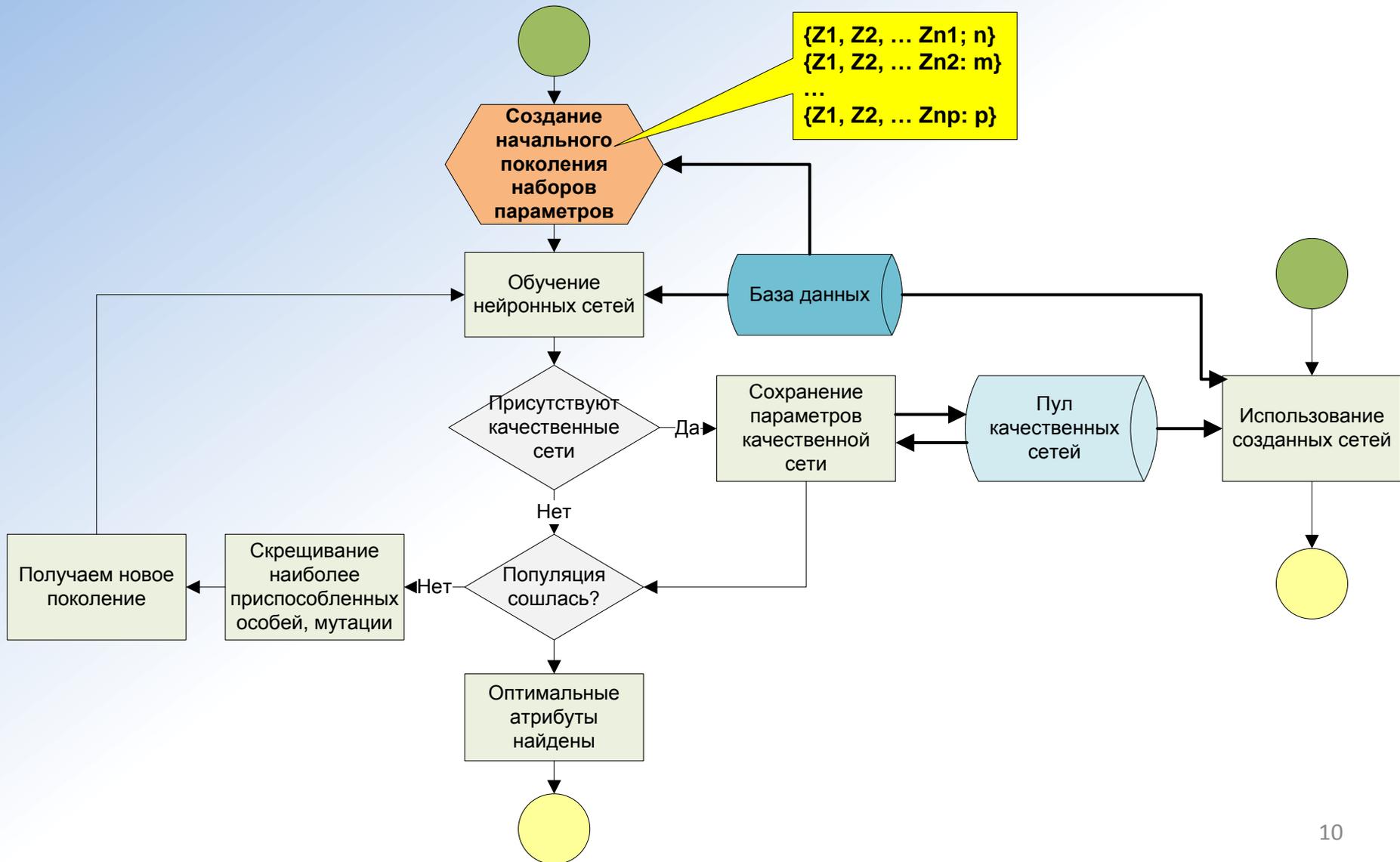


- Замена нечетких переменных интервалами и финальных интервалов нечеткими переменными

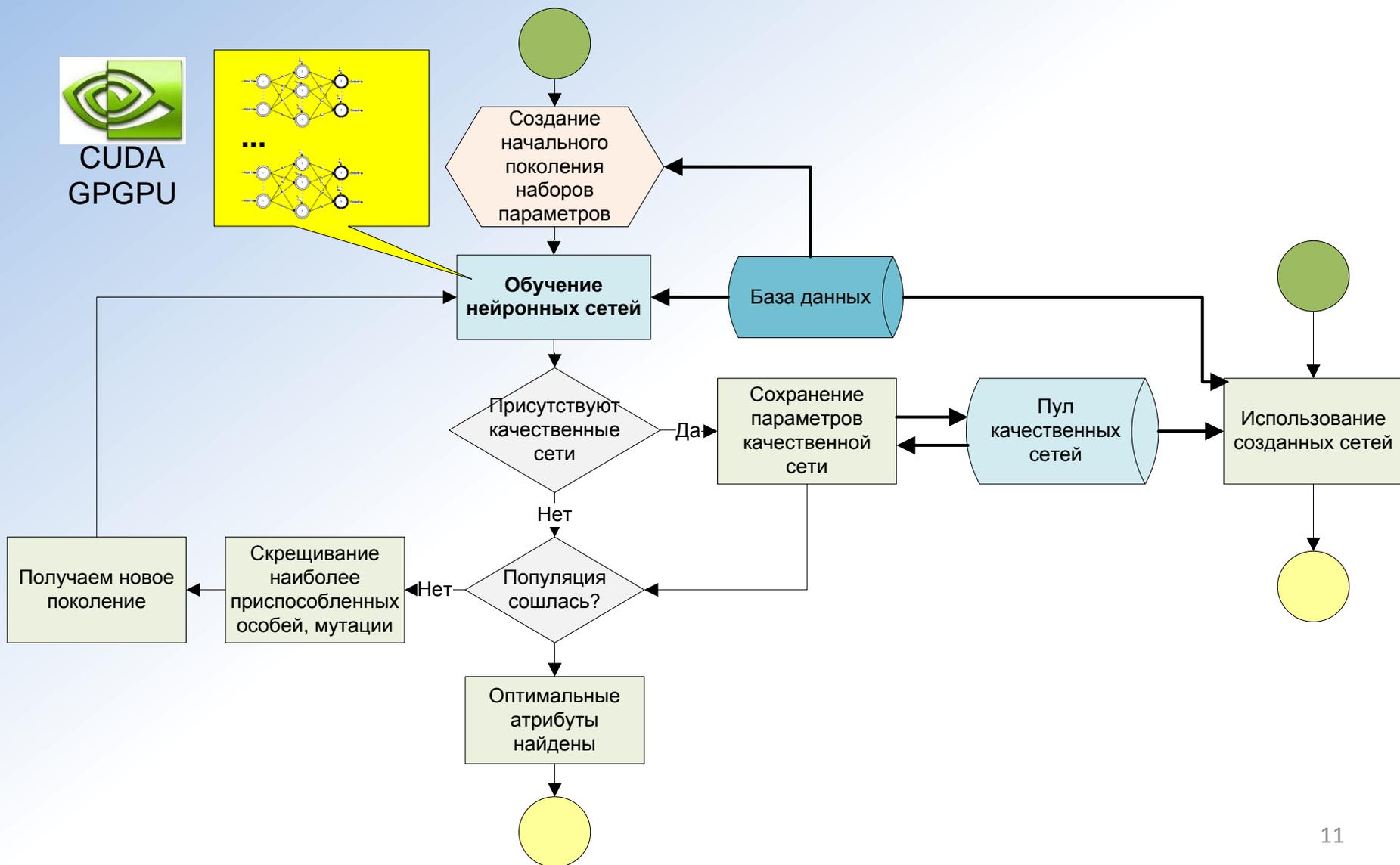
Разработанная система



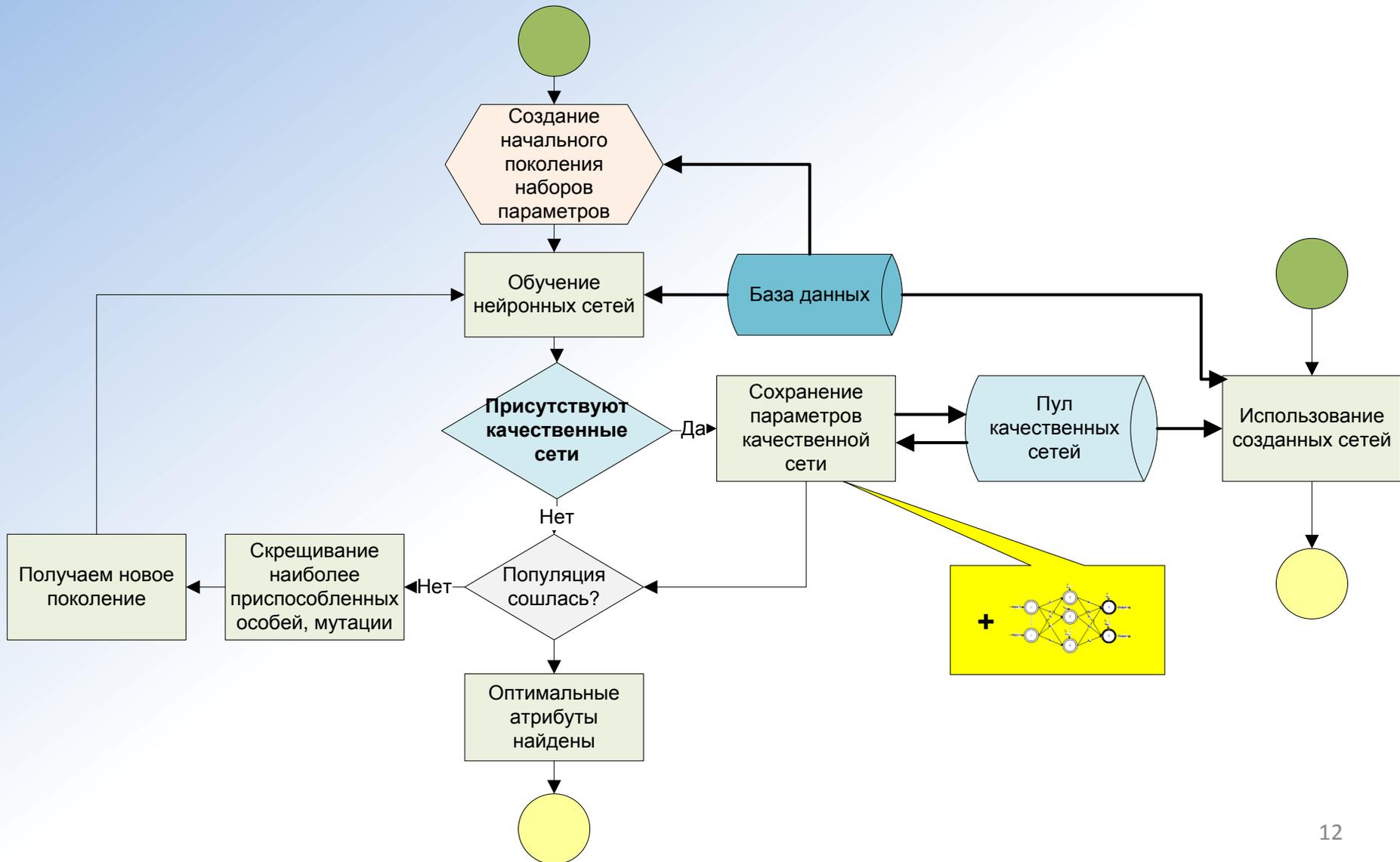
Разработанная система



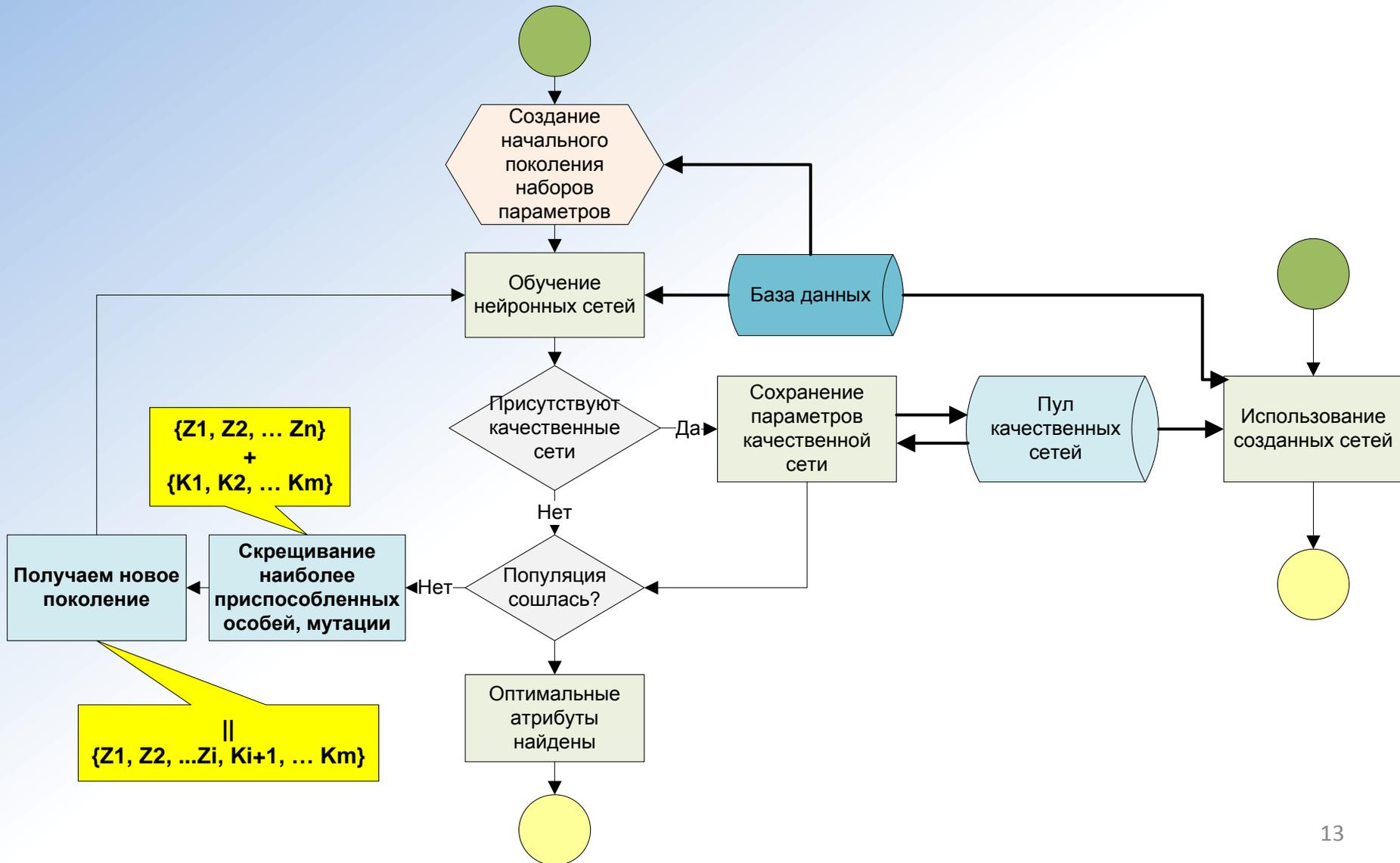
Разработанная система



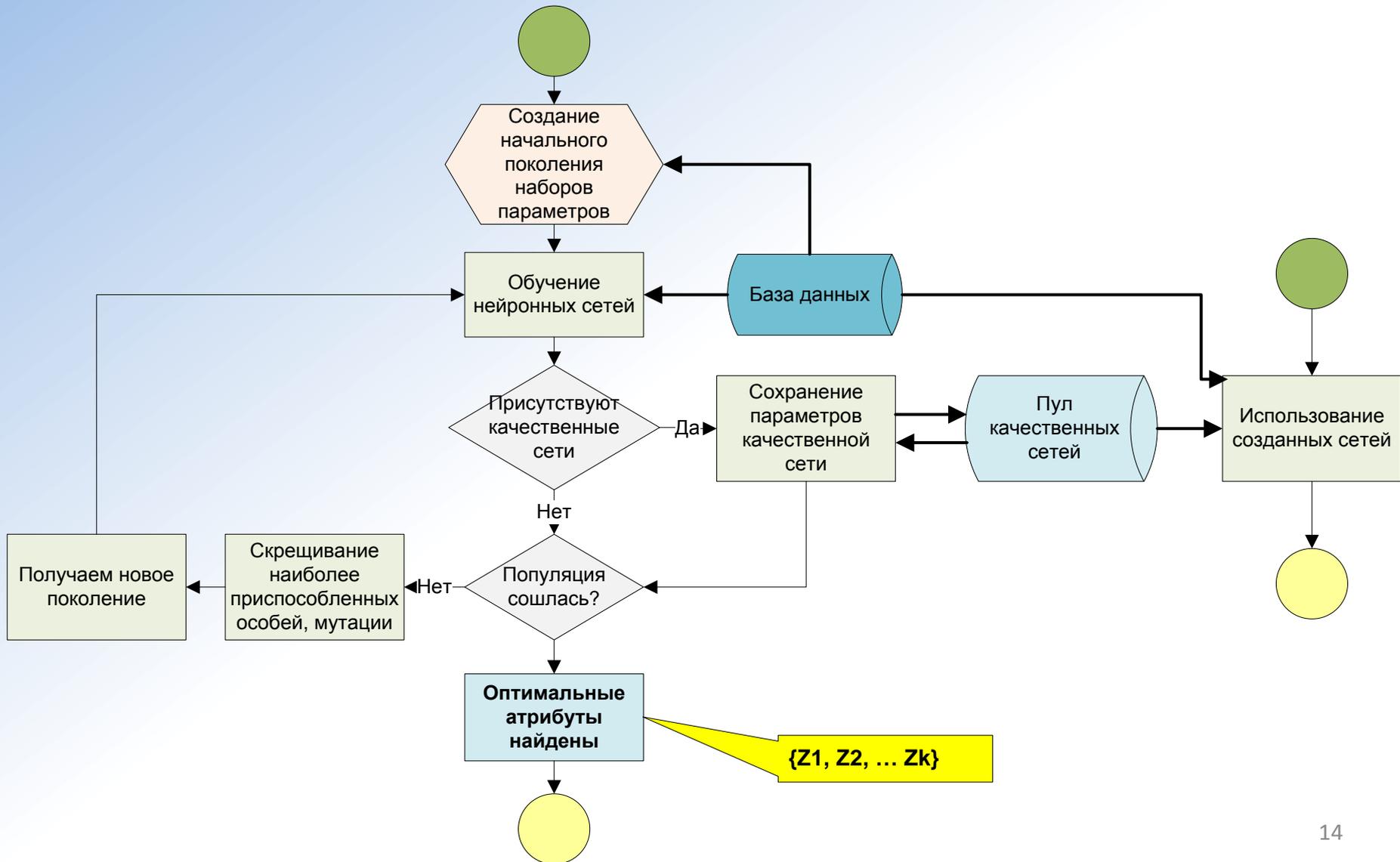
Разработанная система



Разработанная система

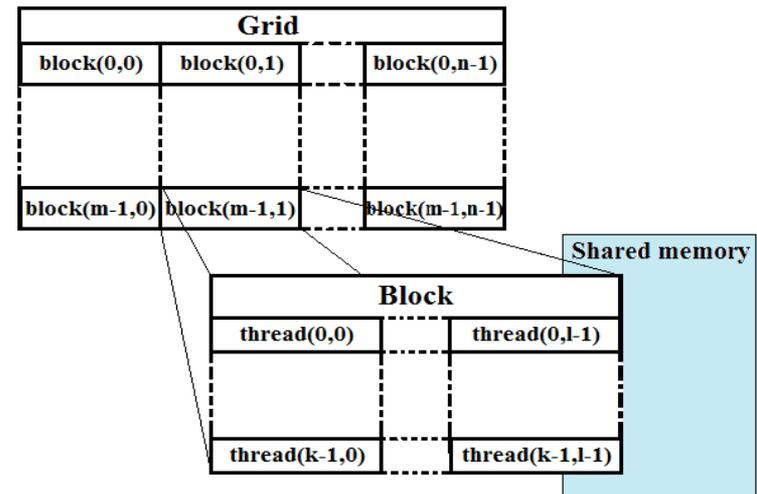


Разработанная система



Архитектура CUDA

- Массивно-параллельные вычисления на ядрах графического процессора
- Совместное параллельное использование CPU и GPU
- Взаимодействие и синхронизация нитей
- CUDA C



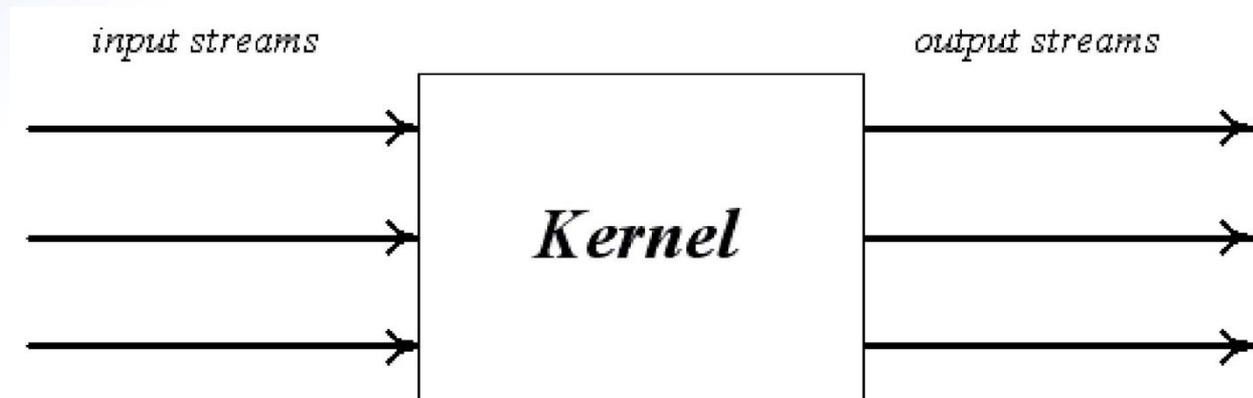
Зачем нужны многоядерные системы

- Минимальный рост частот
 - Энерговыведение \sim четвертой степени частоты
 - Ограничения техпроцесса
- Развитие многоядерных систем
- Видно, что одноядерные системы зашли в тупик
- Повышение быстродействия следует ждать именно от параллельности

Виды параллелизма

	Single Instruction	Multiple Instruction
Single Data	SISD	MISD
Multiple Data	SIMD	MIMD

GPU - SIMD



Программная модель CUDA

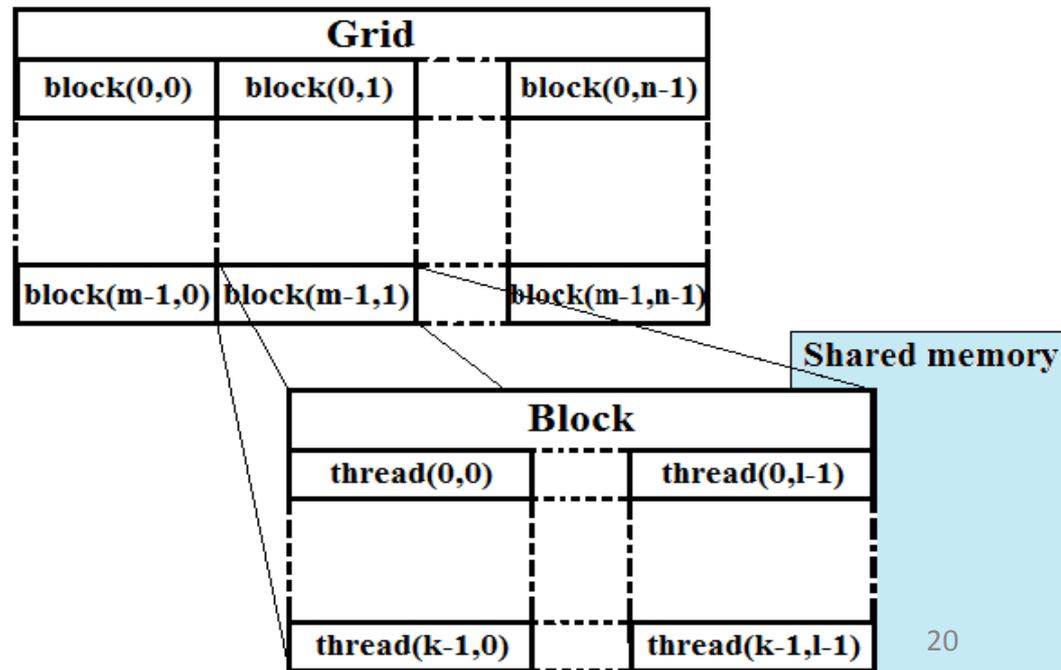
- GPU – это вычислительное устройство, которое:
 - Является сопроцессором к CPU
 - Имеет собственную память (DRAM)
 - Выполняет одновременно сотни тысяч нитей
- Код состоит как из последовательных, так и из параллельных частей
- Последовательные части кода выполняются на CPU
- Массивно-параллельные части кода выполняются на GPU как ядра

Программная модель CUDA

- Отличия нитей между CPU и GPU
 - Нити на GPU очень «легкие»
 - Почти нулевые затраты планировщика задач
 - Для полноценной загрузки GPU нужны тысячи нитей
- Параллельная часть кода выполняется как очень большое количество нитей
- Нити группируются в блоки (blocks) фиксированного размера
- Блоки объединяются в сеть блоков (grid)
- Ядро выполняется на сетке из блоков

Программная модель CUDA

- Каждая нить и блок имеют свой идентификатор
- Нити одного блока могут взаимодействовать между собой
- Нити разным блоков не могут взаимодействовать между собой



Распараллеливание нейронной сети

Перед началом обучения сети

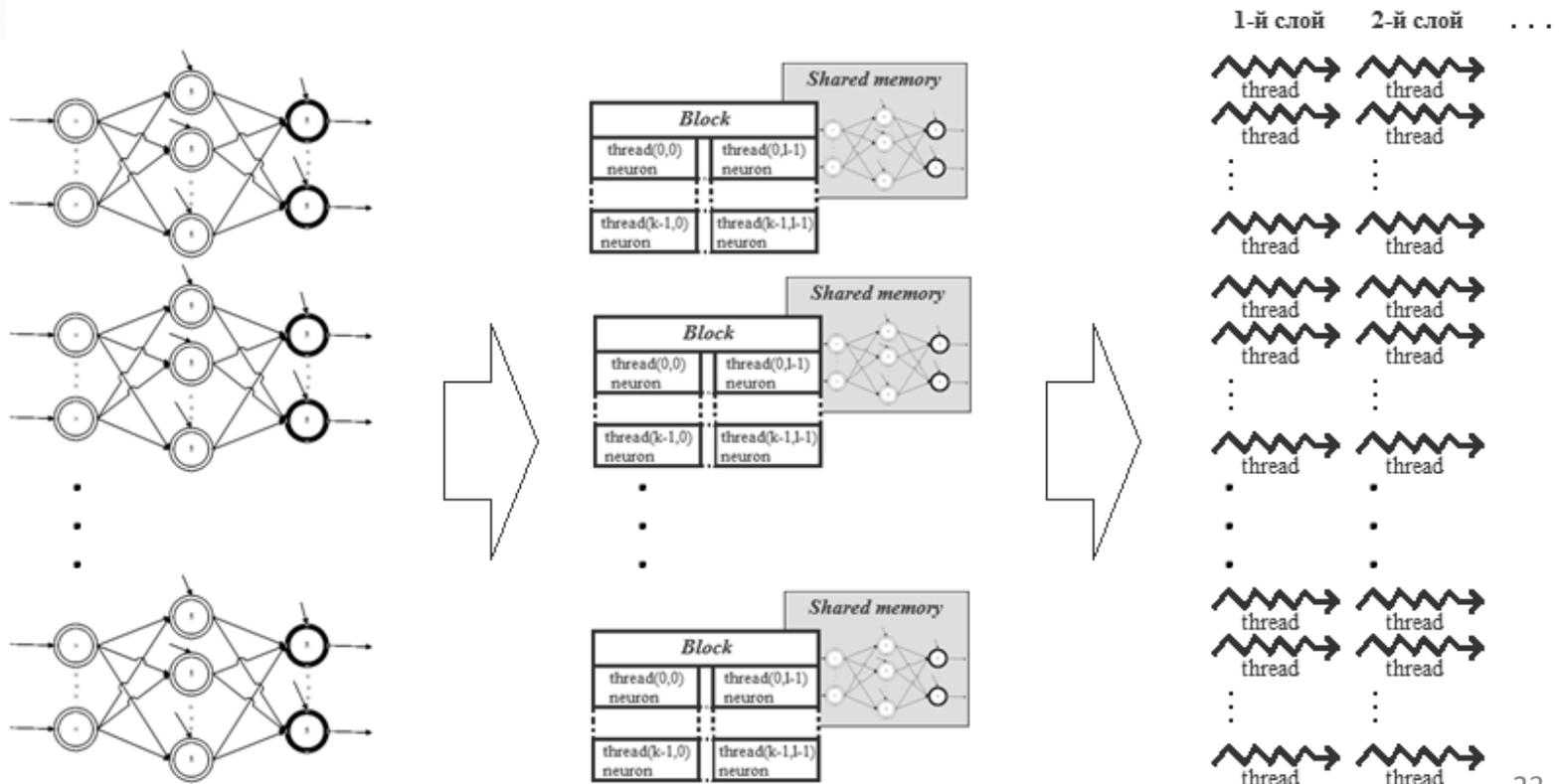
- все параметры сети заносятся в общую память GPU
- Выделяется необходимый объем разделяемой памяти
- Задаются размеры блока нитей: $M \times N$
 - Где M и N – первое и второе по величине количество нейронов в слое сети

Распараллеливание нейронной сети

1. Копирование всеми нитями данных из общей памяти GPU в разделяемую память
2. Вычисление каждой нитью изменение сигнала на каждой синаптической связи между входным и первым скрытым слоем
3. Парное параллельное суммирование сигналов поступающих на нейроны первого скрытого слоя
4. Вычисление функций активации – значения на выходе первого слоя
5. Аналогично с п.2 для остальных слоев
6. Обучение происходит аналогично, в обратном порядке

Распараллеливание генетических алгоритмов

- Распараллеливание генетического алгоритма достигается за счет запуска обучения всех сетей одного поколения в разных потоках (CUDA stream)



Достигнутые показатели по скорости

Тип вычислений	Число сетей	Количество итераций	Количество циклов обучения	Время работы
Последовательные	1	10	1,000	0.51 sec.
	2	10	1,000	1.07 sec.
	3	10	1,000	1.66 sec.
Параллельные (GPGPU)	1	10	1,000	0.47 sec.
	2	10	1,000	0.53 sec.
	3	10	1,000	0.58 sec.

Использовались одинаковые наборы сетей, Каждая сеть в наборе состояла из 10 входных, 15 скрытых и 2 выходных нейронов

Достигнутые результаты

- Разработан алгоритм обучения интервальной нейронной сети (обобщение алгоритма обратного распространения ошибки)
- Создана принципиально новая модель оболочки СППР
- Разработан алгоритм параллельной обработки созданной системы
- Разработана программа, реализующая разработанную оболочку СППР.

Научная новизна

- Предложен принципиально новый метод нахождения оптимальной структуры нейронной сети

Практическая значимость

- Разработан программный продукт
- Повышено качество обучения нейронных сетей
- Увеличена скорость обучения наборов нейронных сетей

Спасибо за внимание!