

Алгоритмы рисования линейных диаграмм

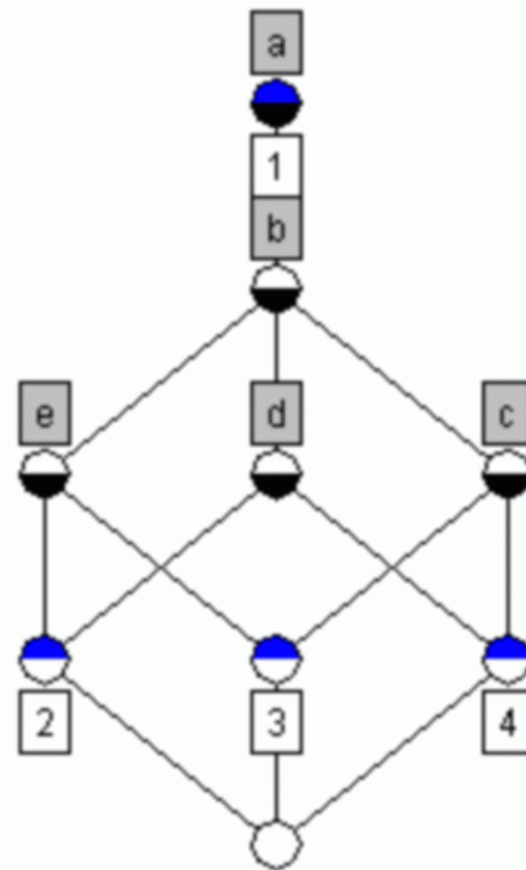
Сергей Евтушенко, НТУУ “КПИ”

Формальный анализ понятий

- Логико-алгебраический метод интеллектуального анализа данных
- Основа – возможность однозначного построения по произвольному бинарному отношению полной решетки
- При анализе не происходит потери информации
- Теоретическая база нескольких методов интеллектуального анализа данных

Формальный контекст и линейная диаграмма

	a	b	c	d	e
1	×				
2	×	×		×	×
3	×	×	×		×
4	×	×	×	×	



Общие требования, предъявляемые при рисовании графов

- две вершины не должны быть расположены в одной точке;
- количество пересечений между ребрами должно быть как можно меньше;
- ребро не должно пересекать вершину, не являющуюся ее концом

Требования, специфические для линейных диаграмм

- Все подпонятия какого-либо понятия должны находиться ниже его;
- ребра изображаются в виде прямых линий;
- наглядное представление структуры решетки;
- использование как можно меньшего количества различных направлений;
- максимизация количества параллельных линий.

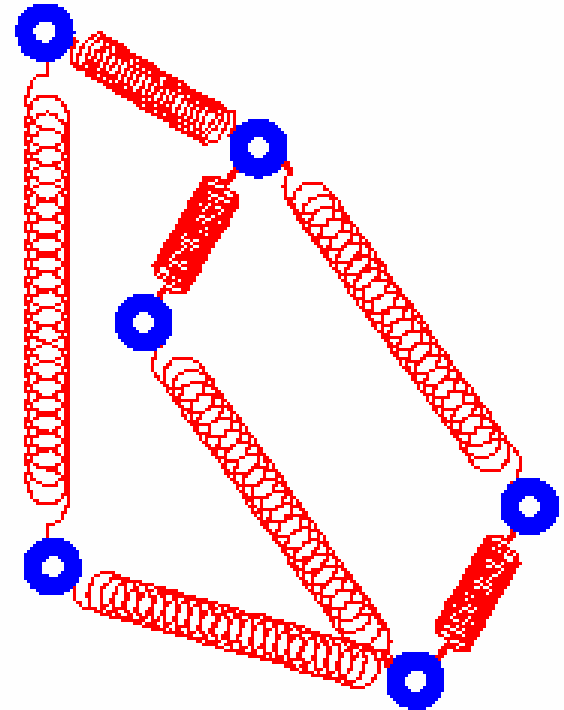
Классические методы рисования ориентированных графов

1. преобразование графа в двухслойный
2. минимизация пересечений между соседними слоями при помощи некоторой эвристики;
3. расчет положений вершин;
4. непосредственное рисование графа.

Задача минимизации пересечений уже для двухслойного графа является NP - трудной

Методы взаимодействия сил (Force-Directed Layout)

- Задается система сил притяжения и отталкивания, действующих между вершинами и ребрами графа
- Находится положение, в котором система имеет минимум энергии (локальный)
- Можно задавать различные системы сил



Эвристика параллелограмма

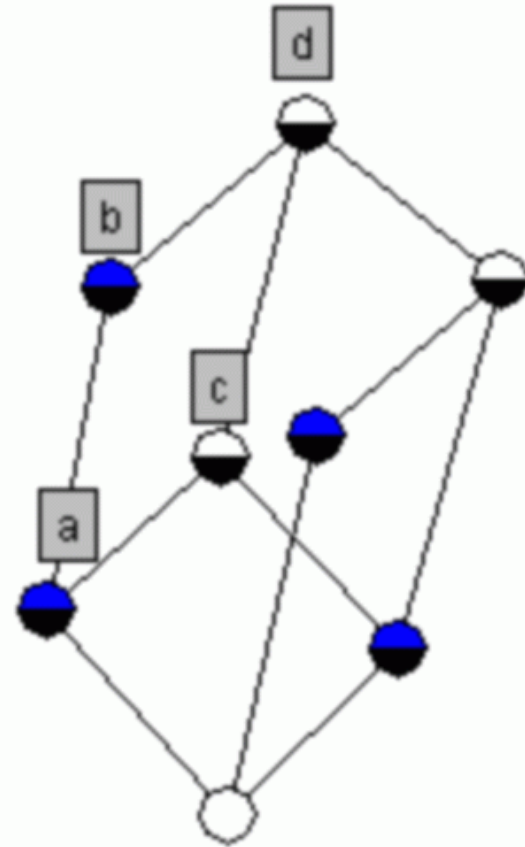
Каждая четверка
элементов

a, b, c, d , таких, что

$$a < b \ \& \ a < c \ \&$$

$$b < d \ \& \ c < d$$

изображается как
параллелограмм.



Аддитивные линейные диаграммы

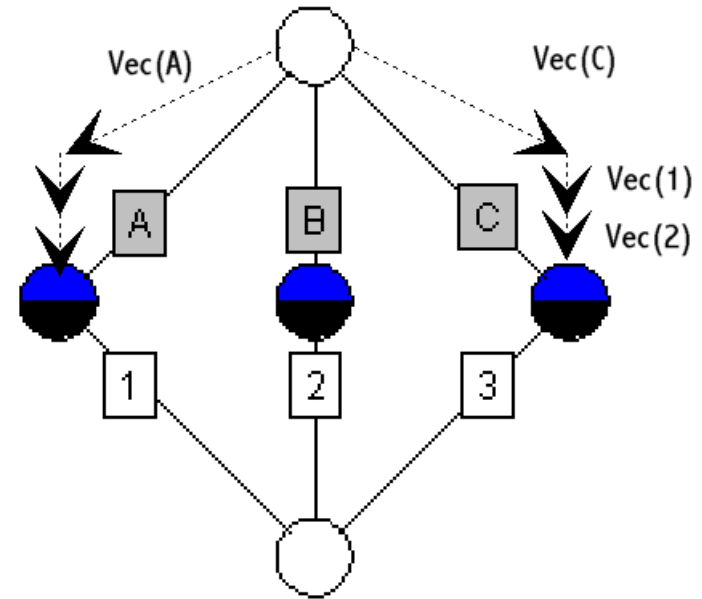
Вводятся

- множество представления R
- $\text{rep}: L \rightarrow \text{Pow}(R)$,
 $x \leq y \Rightarrow \text{rep}(x) \subseteq \text{rep}(y)$

$$\forall r \in R \Rightarrow \text{vec}_r$$

Позиция элемента решетки

$$\text{pos}(x) = n + \sum_{r \in \text{rep}(x)} \text{vec}_r$$



$$G = \{1, 2, 3\} \quad M = \{A, B, C\}$$
$$\text{rep}(\{3\}, \{C\}) = (\{C, 1, 2\})$$

Метод разложения на цепи

- Решетка представляется как произведение цепей
- Для выделения цепей рассчитывается граф отношения порядка между «несводимыми» атрибутами. $m_1 < m_2$, если $\{m_1\}' \subset \{m_2\}'$
- Граф отношения порядка является двудольным. В нем находится максимальное паросочетание, из которого и определяются цепи

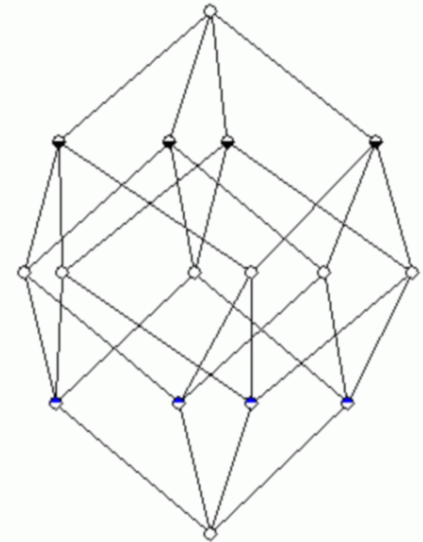
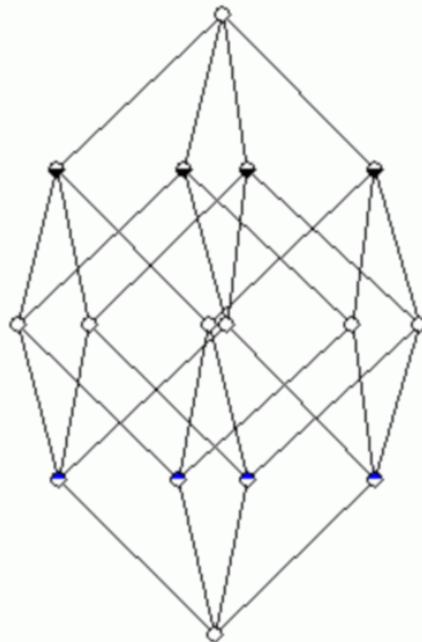
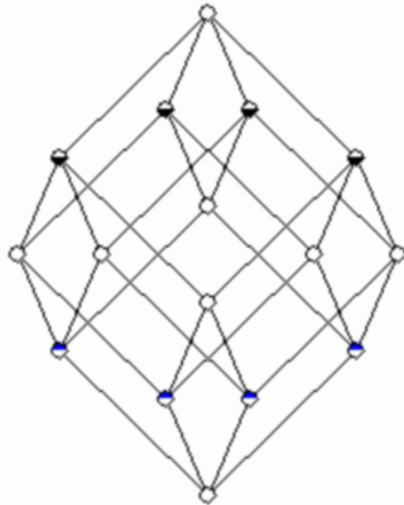
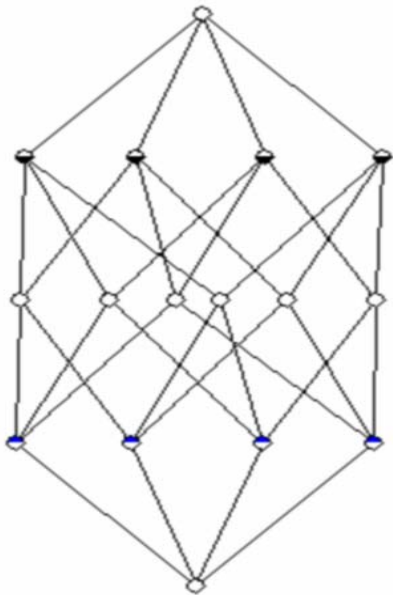
Сравниваемые алгоритмы

Проведено сравнение следующих алгоритмов:

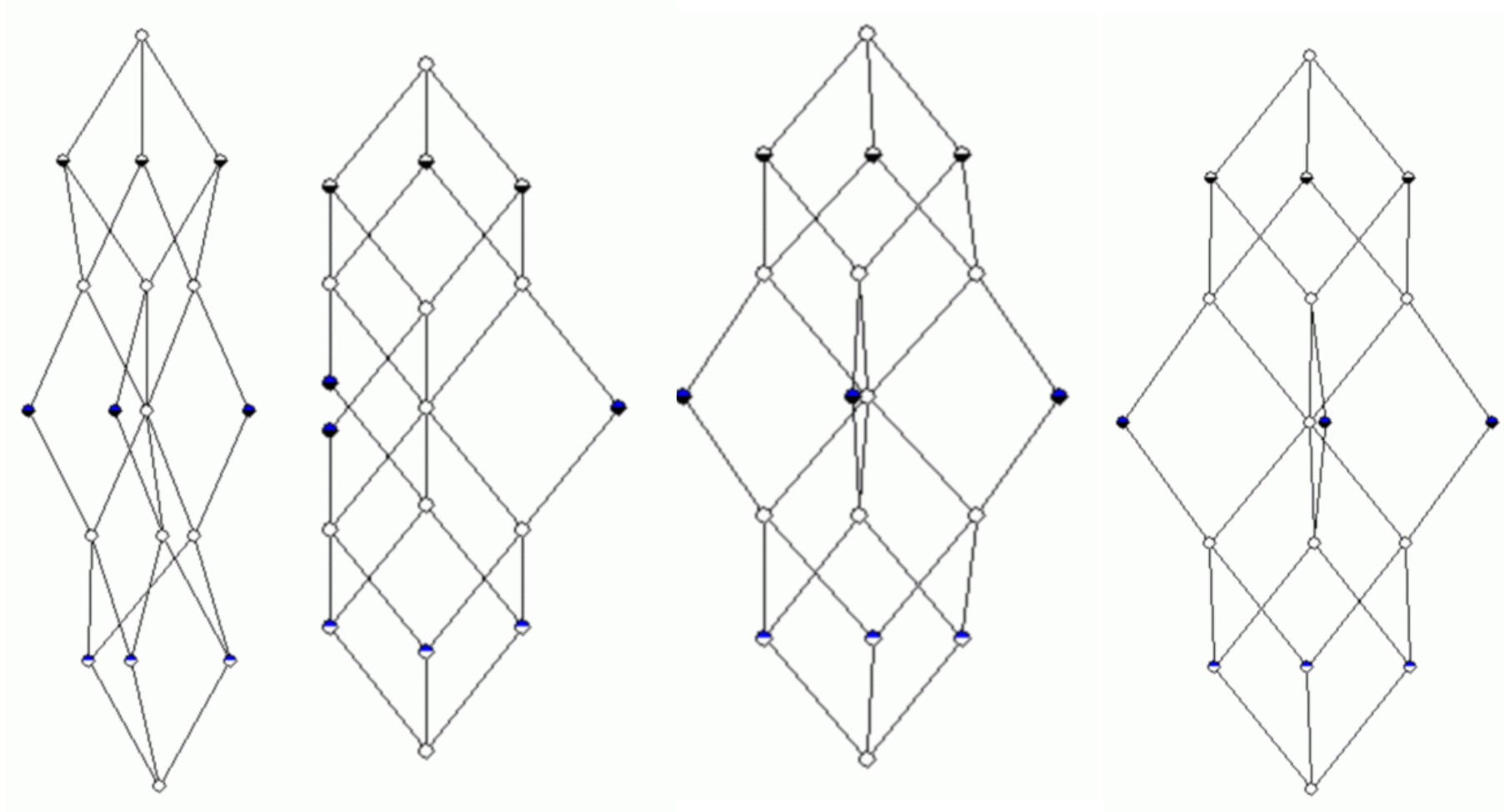
1. Рисования иерархических графов с минимизацией пересечений
2. Разложения решетки на цепи
3. Двух методов взаимодействия сил с использованием разных моделей сил

Для сравнения использовались контексты, которые приводились в публикациях по формальному анализу понятий

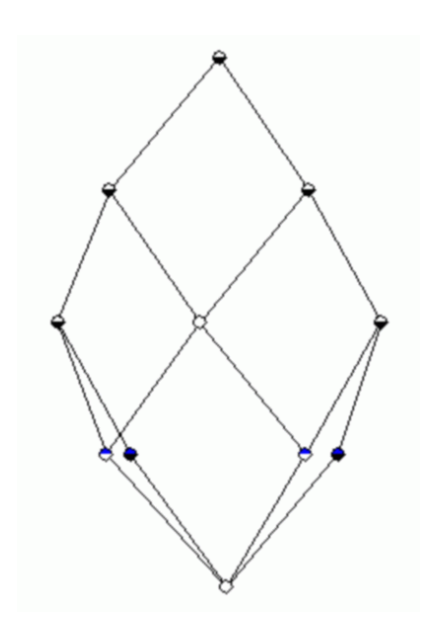
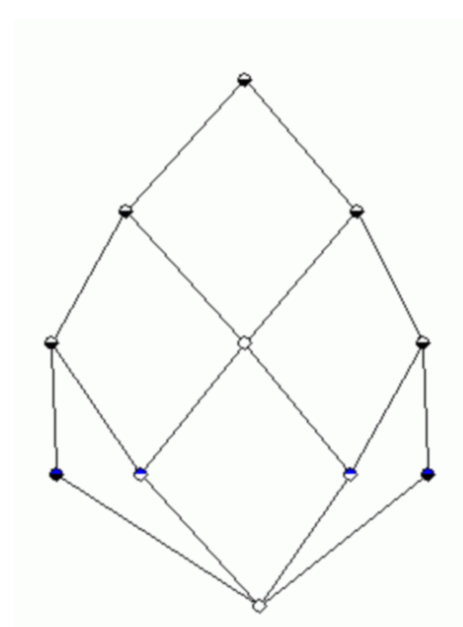
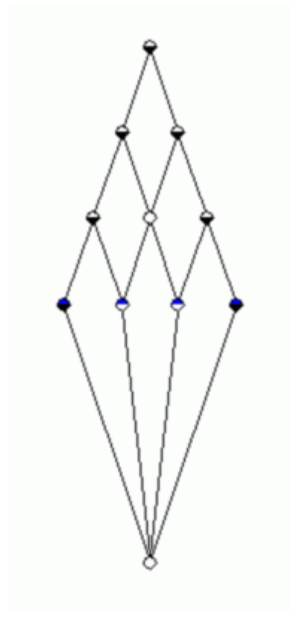
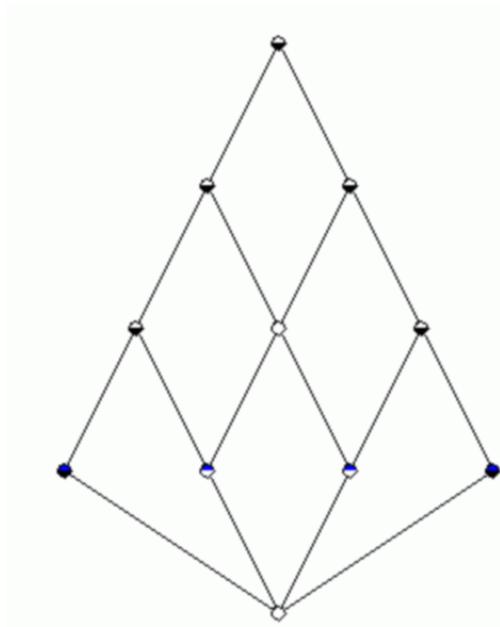
B4



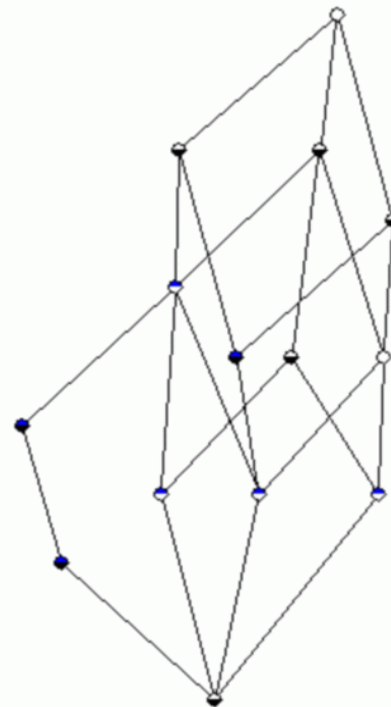
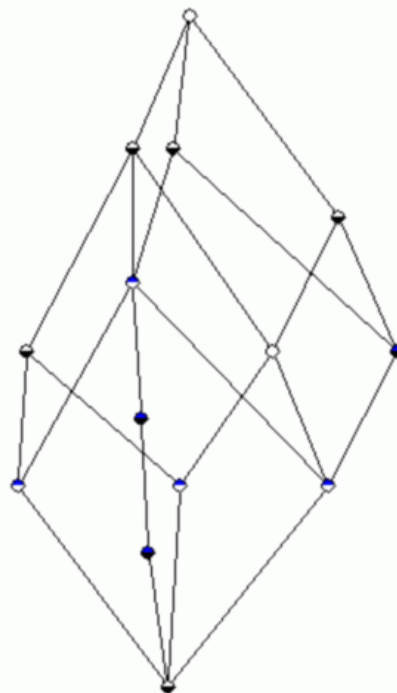
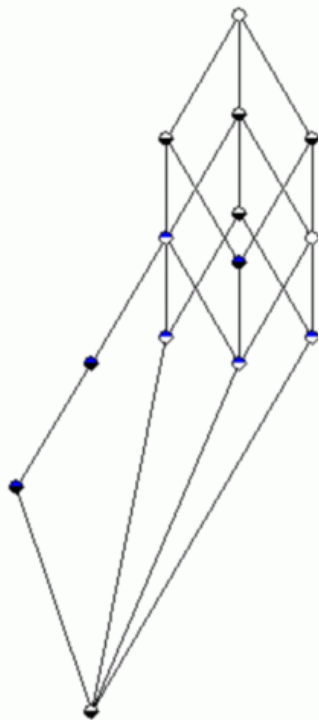
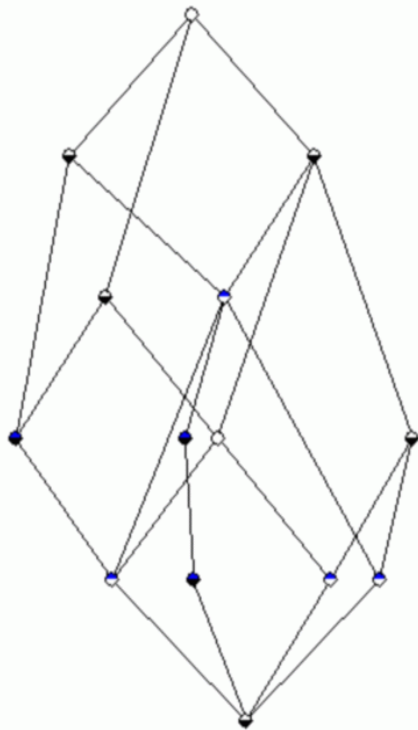
FD3



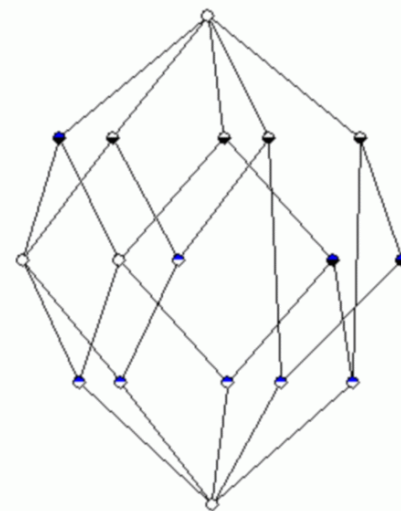
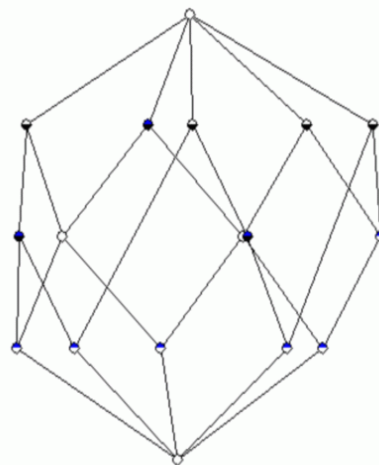
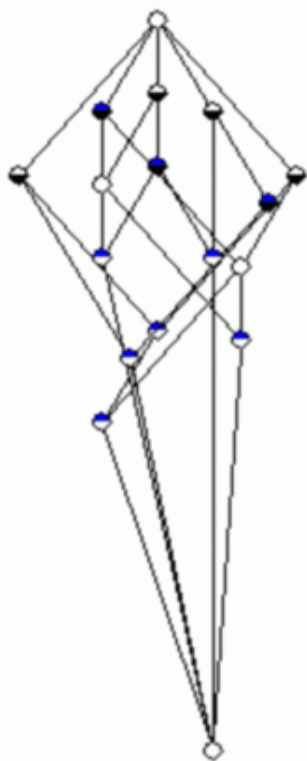
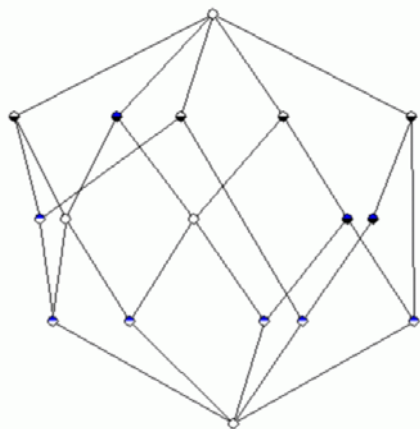
ID4



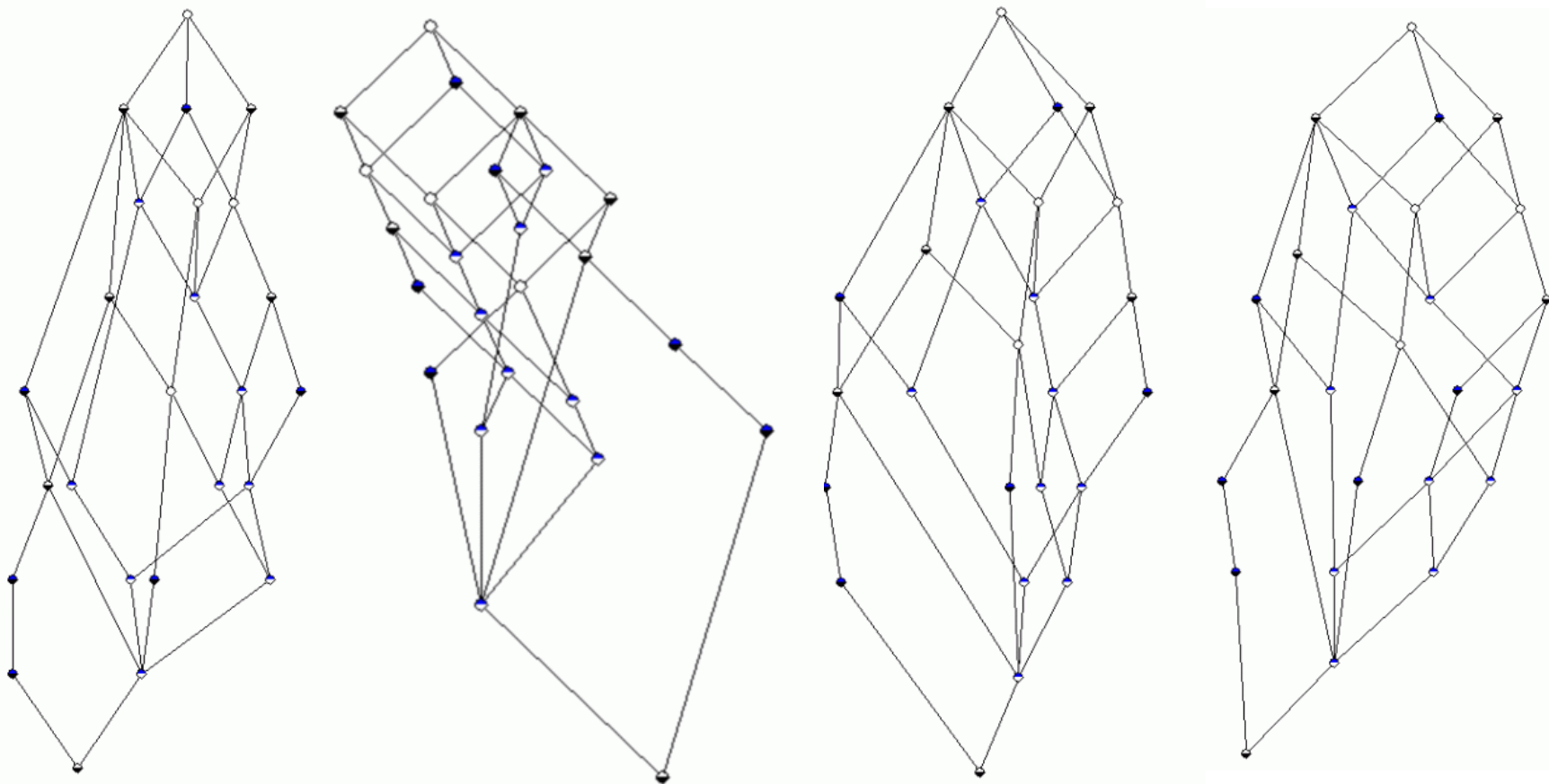
Отношение студента к предметам



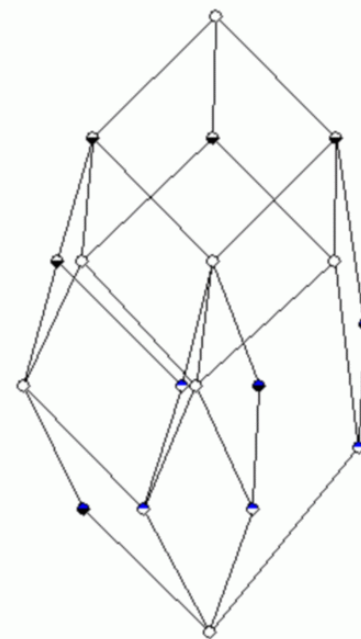
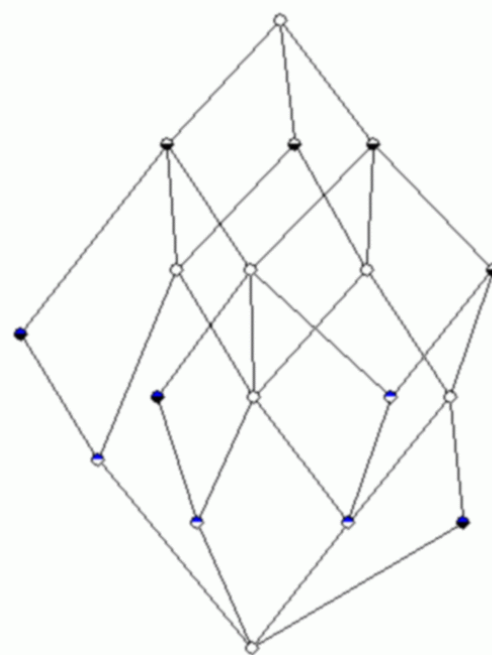
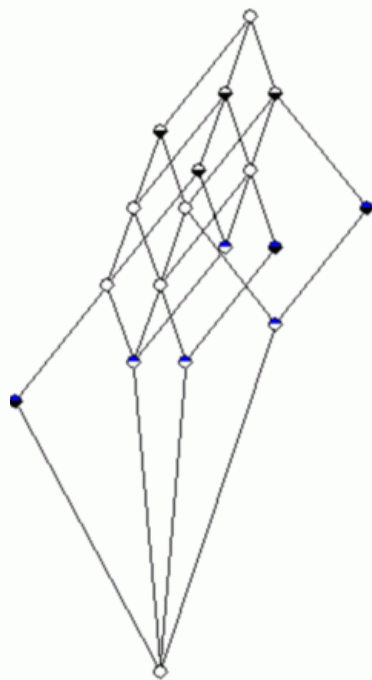
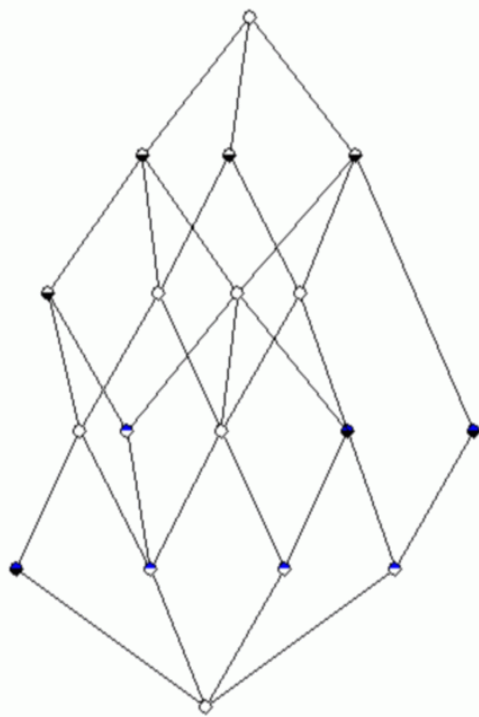
Цвета



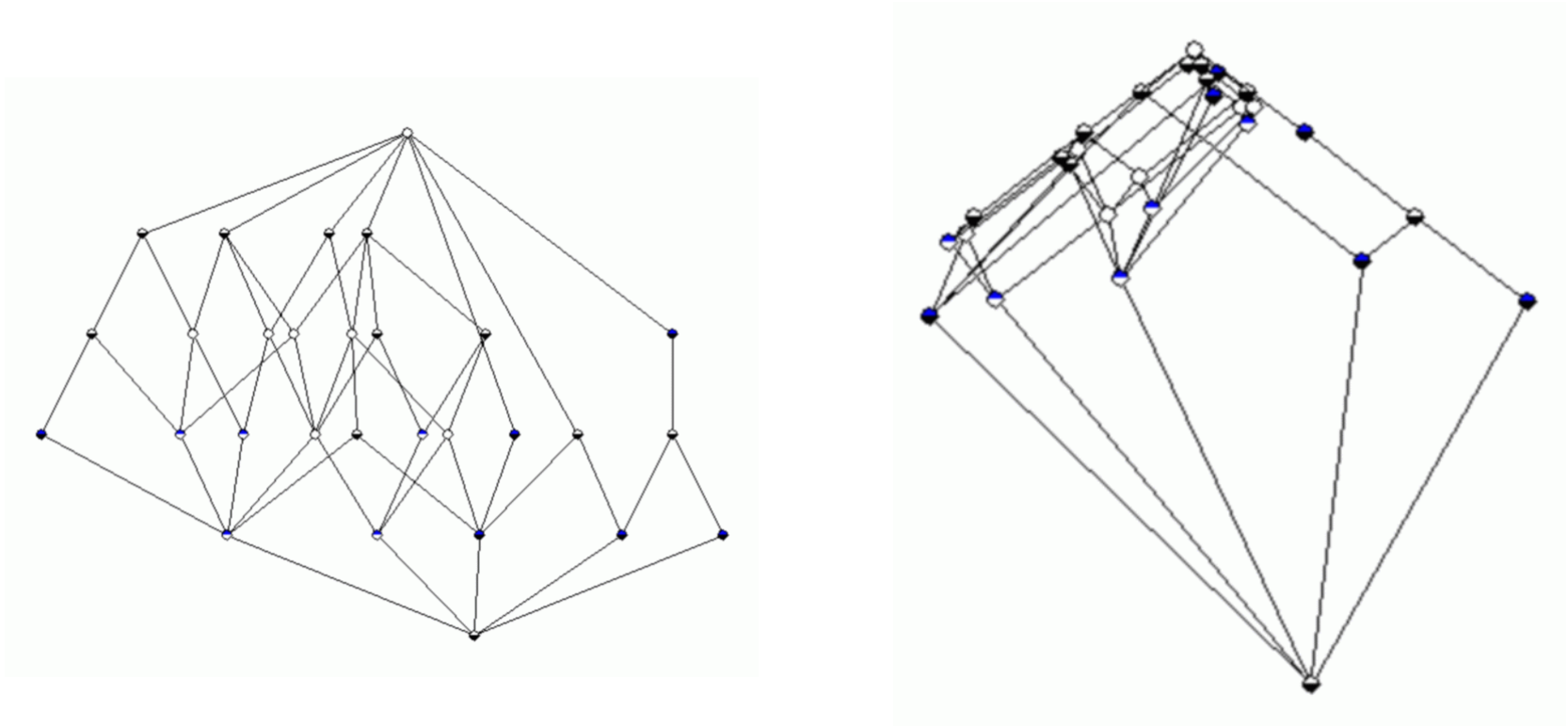
Свойства конечных решеток



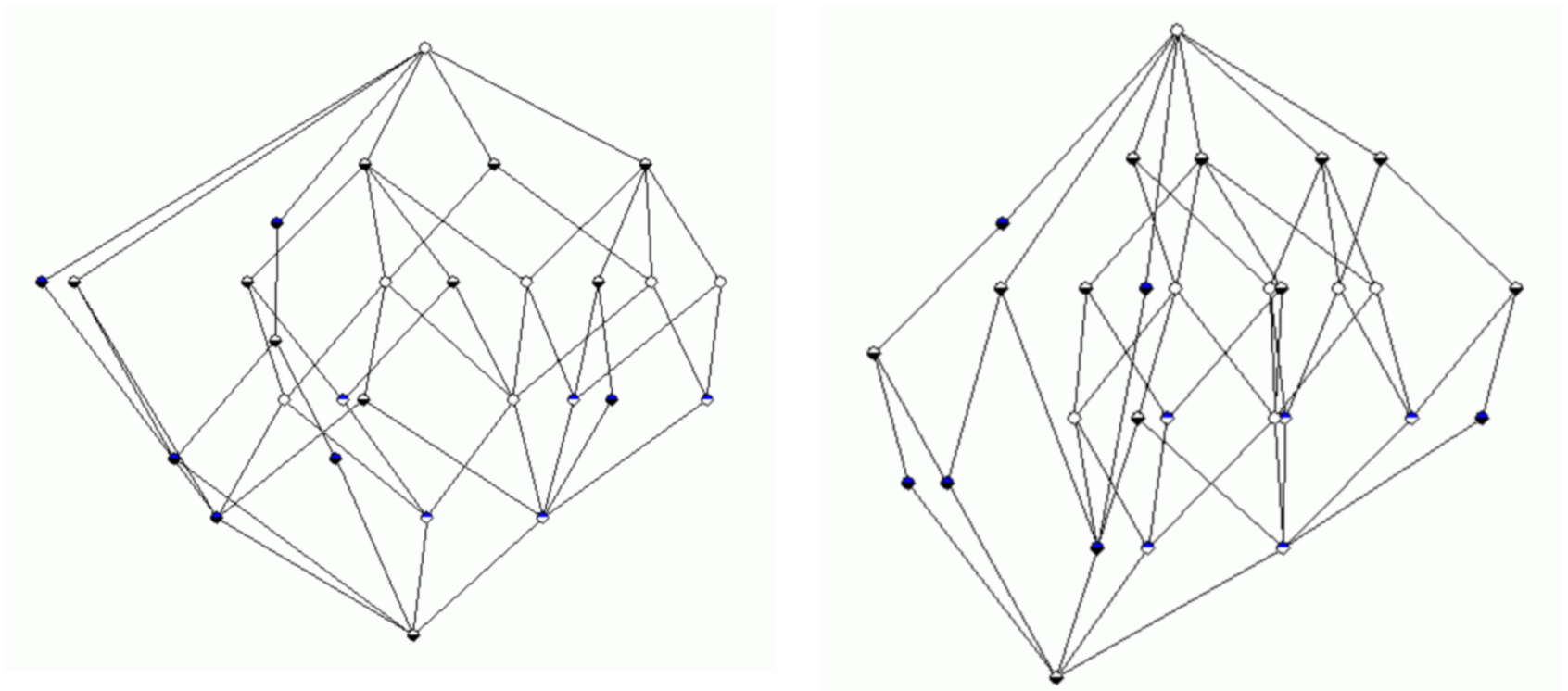
Треугольники



Репертуарная решетка оценки БОЛЬНЫМ СВОИХ ЗНАКОМЫХ (1)



Репертуарная решетка оценки БОЛЬНЫМ СВОИХ ЗНАКОМЫХ (2)



Метод минимизации пересечений

Достоинства:

высокая наглядность изображения

Недостатки:

не выявляет симметричные части

невысокое быстродействие

Метод разложения на цепи

Достоинства

- Высокое быстродействие
- Максимизация параллельности линий

Недостатки

- Плохо работает для недистрибутивных решеток
- Выбор стандартных направляющих векторов не гарантирует отсутствия перекрытия разных вершин
- Проблема «отделенного нулевого элемента»

Методы взаимодействия сил

Достоинства

- Выявляют симметрии
- Работают в 3-х мерном пространстве

Недостатки

- Низкое быстродействие
- Не всегда сходятся к оптимальному решению
- Не обеспечивают максимизацию параллельности линий