

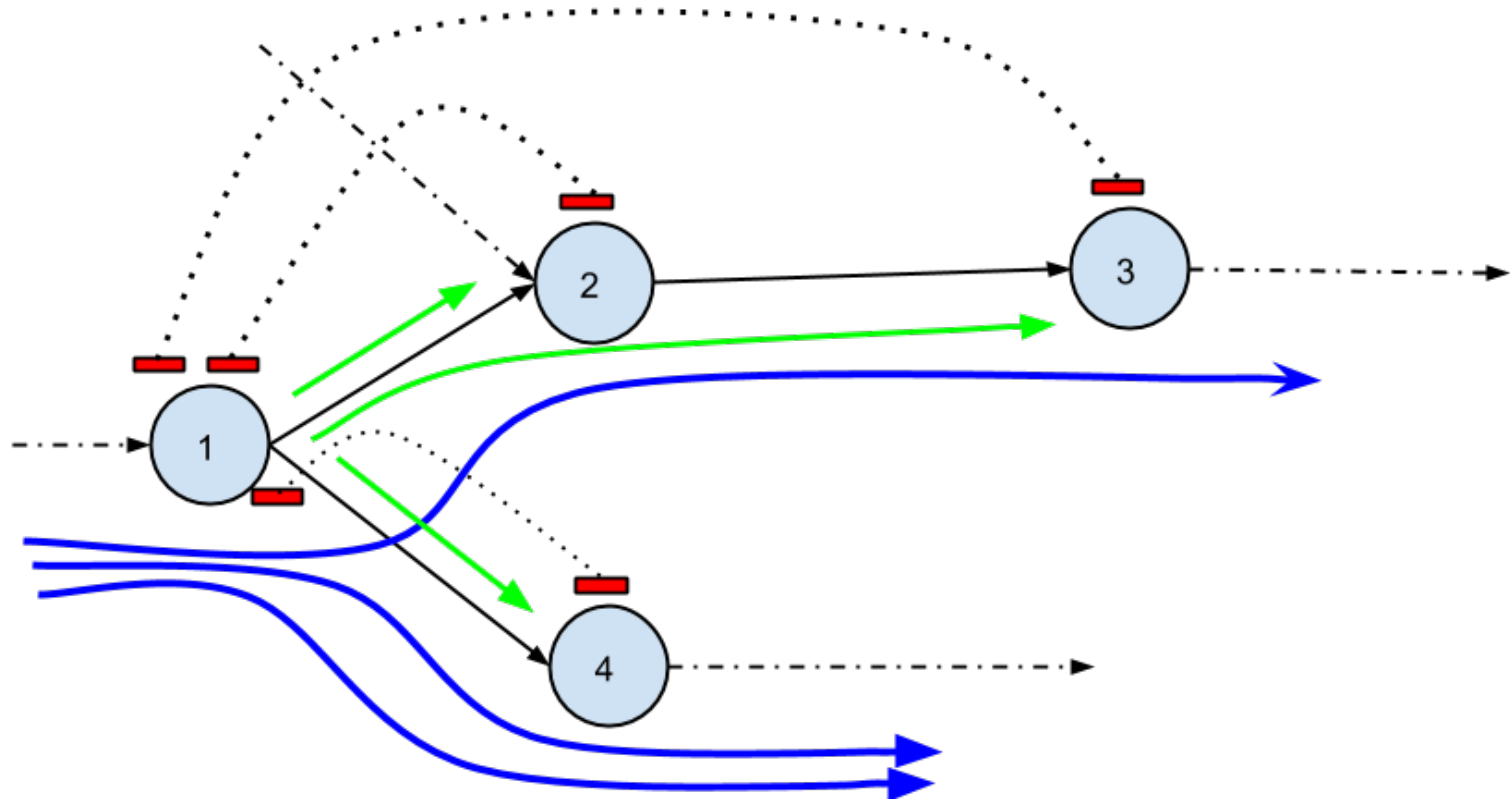


Определение копийности контигов больших геномов

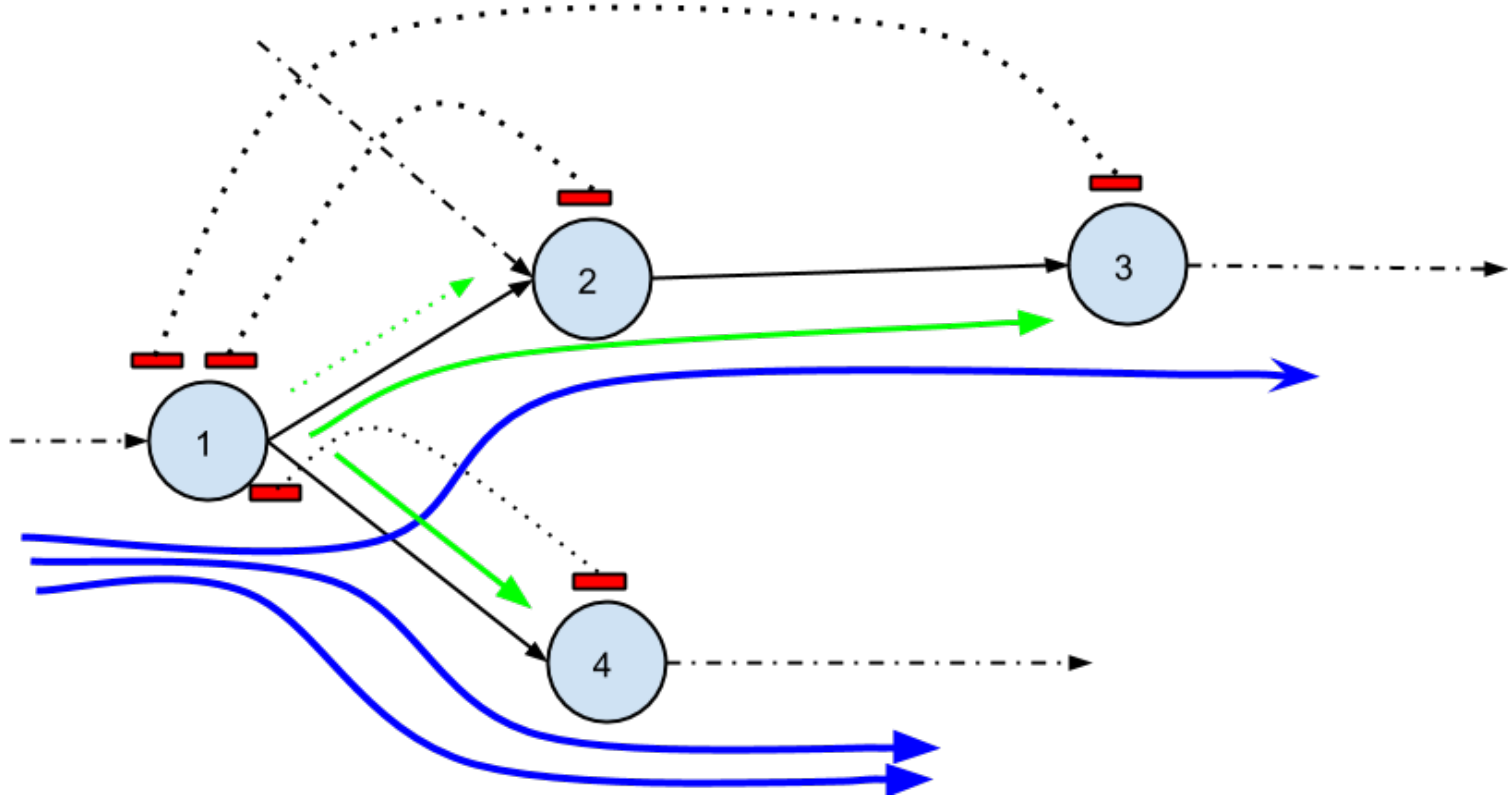
Бушманова Вера (СПбГУ)

Руководители: Мелешко Дима (АУ), Son Pham (UCSD)

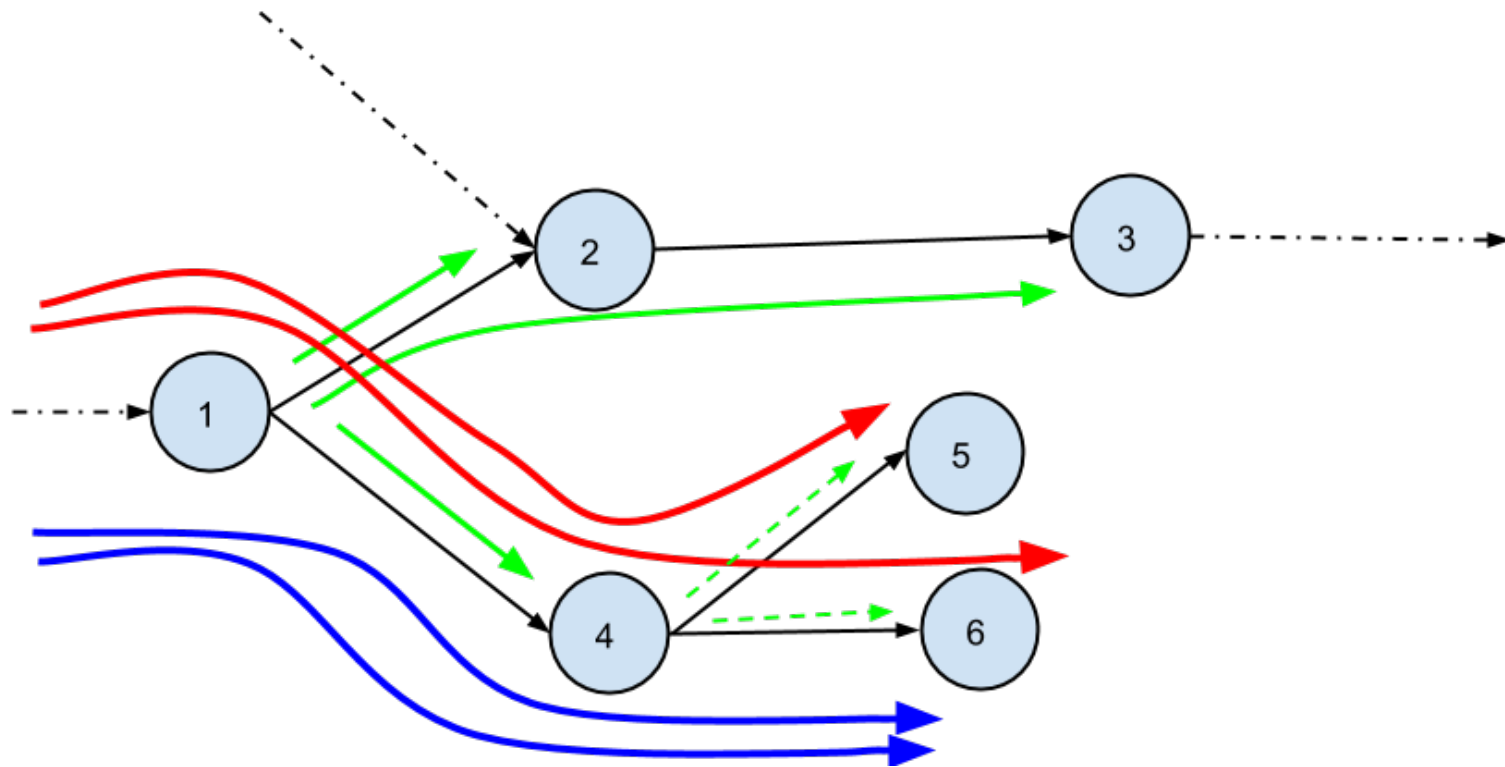
Шаг1.Находим правильные пути по информации о контигах



Шаг2. Удаляем вложенные пути



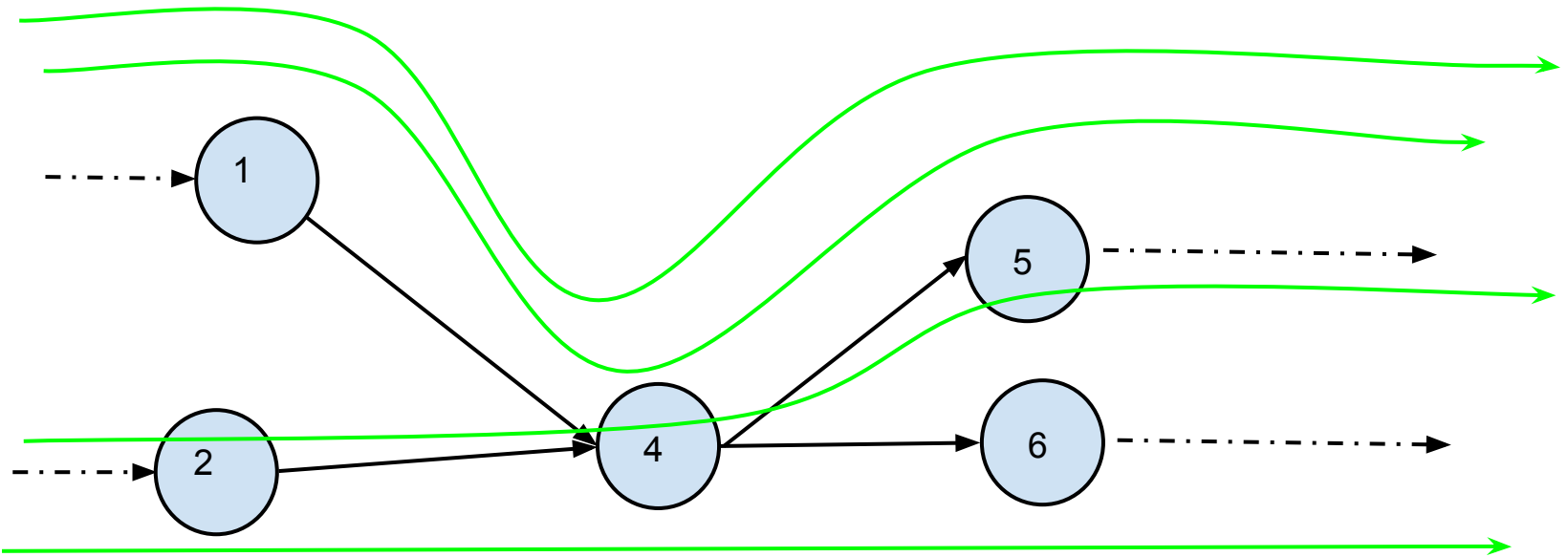
Шаг 3. Удлиняем пути--предсказываем CN



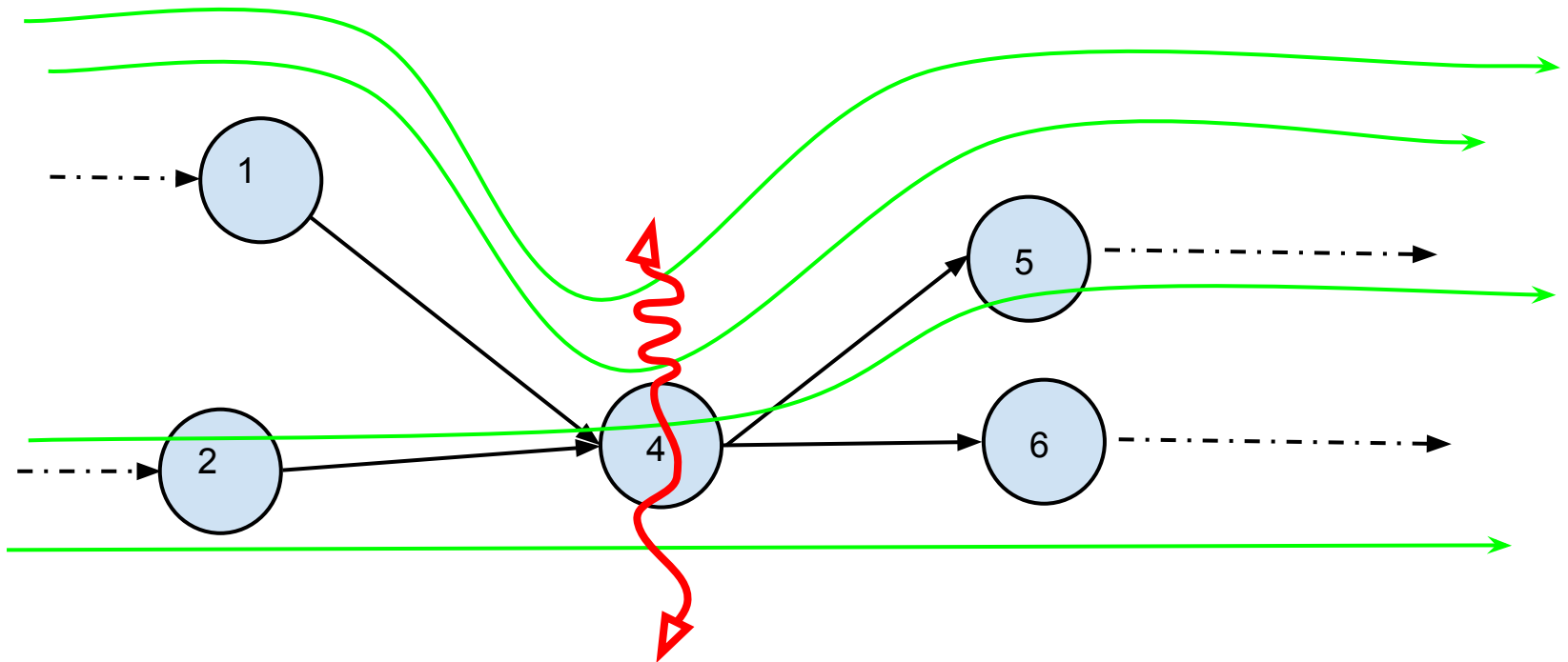
Проблема

Для коротких контигов нет парной информации->путей

Решение



Решение



Повторяем шаги 1-3 для вершины 3

Метрики

КОНТИГИ:

короткие <250bp

средние 250-1000bp

длинные >1000bp

1) Правильно предсказанные CN

2) Сумма ошибок = Сумма $|CN_{\text{true}} - CN_{\text{predicted}}|$

Результаты E.Coli

	Правильно предсказанные CN (короткие/средние/ длинные/все контиги)	Сумма ошибок (короткие/средние/ длинные/все контиги)
Magnolya	49/14/122/185	751/58/3/812
CNVera (старая версия)	93/15/111/219	424/46/20/490
CNVera (новая версия)	106/13/113/232	366/44/16/426

В сборке всего 281 коротких, 43 средних и 125 длинных контигов, всего - 449

Результаты *s. cerevisiae*

	Правильно предсказанные CN (короткие/средние/ длинные/все контиги)	Сумма ошибок (короткие/средние/ длинные/все контиги)
Magnolya	346/143/302/791	18341/4832/17/2319 1
CNVera (старая версия)	383/138/286/807	17084/4650/30/2176 4
CNVera (новая версия)	355/137/286/778	15402/4546/30/1997 9

В сборке всего 1618 коротких, 580 средних и 312 длинных контигов, всего - 2210

Спасибо за внимание!