Alternative Procedure

Idea: Do not use a single cycle to cover all vertices but a set of cycles
$\rightsquigarrow$ cycle-cover for which each vertex must be part of exactly one cycle.

Here: Cycle-cover of minimal cost (according to distance graph).

Remark: A minimal cycle-cover can be computed in time $\mathcal{O}(n^3)$ (whereas the restriction to edges to be covered by at least one cycle leads to an $\mathcal{NP}$-complete problem).

Approxiamtion: Construct a single cycle (solution to TSP) from the minimal cycle-cover.

Steps:

1. Identify each cycle $c \in \mathcal{C}$ by one of its vertices $\rightsquigarrow$ set of vertices R.
2. Compute a minimal cycle-cover for the subgraph induced by $R \rightsquigarrow$ set of cycles $\mathcal{C}'$.
3. For each $c \in \mathcal{C}'$ delete the edge corresponding to a merge of minimal overlap $\rightsquigarrow$ broken cycles are considered as merged strings.
4. Concatenate these strings and expand each of its elements by the cycle from $\mathcal{C}$ it represents.

$$K_1 = a_1 \to a_2 \to a_3 \to a_4 \to a_5 \to a_1$$

$$K_2 = b_1 \to b_2 \to b_3 \to b_4 \to b_5 \to b_6 \to b_1$$

$$K_3 = c_1 \to c_2 \to c_3 \to c_4 \to c_1$$

$$K_4 = d_1 \to \dots \to d_5 \to d_1$$

$$K_5 = e_1 \to \dots \to e_5 \to e_1$$

$$K_6 = f_1 \to f_2 \to f_1$$
$$K_7 = g_1 \to g_2 \to g_3 \to g_1$$

$$S = \{a_1, \dots, a_4, b_1, \dots, b_6, \dots\}$$

$$K_1' = a_1 \to b_1 \to c_1 \to a_1 \rightsquigarrow \overbrace{\langle b_1, c_1, a_1 \rangle}^{u_1}$$
$$K_2' = d_1 \to e_1 \to d_1 \rightsquigarrow \overbrace{\langle e_1, d_1 \rangle}^{u_2}$$
$$K_3' = f_1 \to g_1 \to f_1 \rightsquigarrow \underbrace{\langle f_1, g_1 \rangle}_{u_3}$$

$$\Longrightarrow w' = u_1 u_2 u_3$$

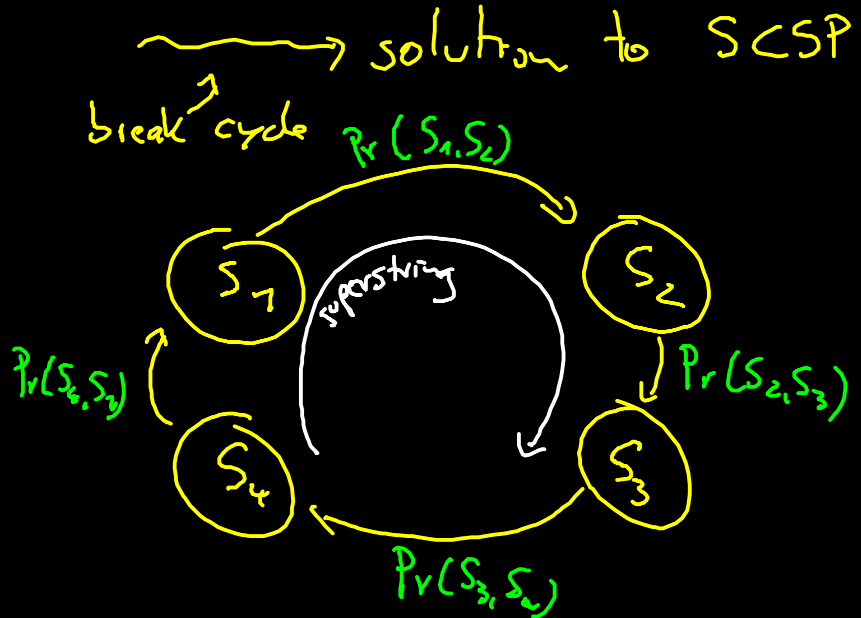
expand

$$w = \langle\langle b_1, b_2, \dots, b_6, b_1 \rangle, \langle c_1, c_2, \dots, c_4, c_1 \rangle, \langle a_1, a_2, \dots, a_4, a_1 \rangle\rangle$$
$$\cdot \langle\langle e_1, \dots, e_5, e_1 \rangle, \langle d_1, \dots, d_5, d_1 \rangle\rangle$$
$$\cdot \langle\langle f_1, f_2, f_1 \rangle \langle g_1, g_2, g_3, g_1 \rangle\rangle$$

Theorem

The algorithm just outlined is a 3-approximation algorithm for SCSP.

Proof: Solution to TSP on digraph



cost for TSP: $pr(S_1,S_2) + pr(S_2,S_3) + pr(S_3,S_4) + pr(S_4,S_1)$ $\rightsquigarrow$ $pr(S_4,S_1) + Pr(S_1,S_2) + Pr(S_2,S_3) + |S_3|$

Since $pr(S_i,S_j) < |S_i| \implies$ cost of TSP lower bound for cost SCSP.

Minimal cycle cover has cost at most the same as TSP $\rightsquigarrow$ min. cycle cover lower bound for TSP.

$$\implies cost(\mathcal{C}) = Opt_{cc}(DG(S)) \leq OPT_{scs}(S)$$

$$cost(\mathcal{C}') = Opt_{cc}(DR(R)) \leq OPT_{scs}(R)$$

since $DG(R)$ is a subgraph of $DG(S)$

$$\Longrightarrow \quad cost(e') \leq cost(e) \qquad (o)$$

minimal overlap of cycle K_i'

Observation: $|U_i| = cost(K_i') + minov(K_i')$

$$\Longrightarrow \quad |w'| \leq cost(e') + \sum_{c' \in e'} minov(c') \qquad (*)$$

Replacing within w' all representatives by their cycle:

gives us, e.g.

$$\underbrace{Pref(a_1,a_2)\cdot Pref(a_2,a_3)\cdot Pref(a_3,a_4)\cdot Pref(a_4,a_1)}_{\text{cost of cycle}}\cdot \underline{\underline{a_1}}$$

$\Longrightarrow$ w' „gets longer" by the cost of resp. cycle.

$$\Longrightarrow \quad |w| \leq |w'| + cost(e)$$

$$\overset{(*)}{\leq} cost(e') + \sum_{c' \in e'} minov(c') + cost(e)$$

$$\overset{(o)}{\leq} 2\cdot cost(e) + \sum_{c' \in e'} minov(c')$$

We obviously have

every cycle at least 2 edges

$$\sum_{c' \in \mathcal{C}'} \mathrm{minov}(c') \le \frac{1}{2} \cdot \sum_{c' \in \mathcal{C}'} \sum_{(s_i, s_j) \in c'} \mathrm{ov}(s_i, s_j)$$

$$\Longrightarrow |w| \le 2 \cdot \mathrm{cost}(\mathcal{C}) + \frac{1}{2} \cdot \sum_{c' \in \mathcal{C}'} \sum_{(s_i, s_j) \in c'} \mathrm{ov}(s_i, s_j)$$

Lemma: Let c_1, c_2 be two cycles of a minimal cycle cover $\mathcal{C}$ and $s_1 \in c_1$, $s_2 \in c_2$ to „strings" of those cycles. Then we have

$$\mathrm{ov}(s_1, s_2) < \mathrm{cost}(c_1) + \mathrm{cost}(c_2).$$

$$\overset{\text{Lemma}}{\Longrightarrow} |w| \le 2 \cdot \mathrm{cost}(\mathcal{C}) + \underbrace{\frac{1}{2} \cdot \sum_{c' \in \mathcal{C}'} \sum_{(s_i, s_j) \in c'} \left(\mathrm{cost}_{\mathcal{C}}(\mathrm{cycle}(s_i)) + \mathrm{cost}_{\mathcal{C}}(\mathrm{cycle}(s_j))\right)}_{\mathrm{cost}(\mathcal{C})}$$

$$= 3 \cdot cost(e)$$

$$\leq 3 \cdot Opt_{scs}(S)$$

$\boxed{4}$