**2. Describe the sequence data you plan to use by specifying:**

. *What the sequences are representing and where they come from :*

The sequence will represent the life course of the post-80s generation in Beijing and of their parents.

N= 900

I am going to collect these data with a life matrix during this academic year.

. *Whether there is one sequence per case or multichannel sequences? In case of*

*multichannel sequences, specify to the following points for each channel*.

One sequence per case.

. *The nature of the sequences (Categorical? Chronological? State or event*

*sequences? ...,)*

Categorical, chronological, state and event sequences as well.

. *The alphabet (list of symbols in the sequences).*

Symbols (ordinal variables)

*{ Is there a natural order of the symbols (ordinal variable)?*

Yes

. *Size of the alphabet*

Max. 9

. *Number of sequences.*

Difficult to answer as I don’t have the data yet but I am planning to have around 50 different sequences.

. *Maximum and minimum sequence lengths.*

Max. length : 33 years

Min. lenght : 23 years

. *The kind of knowledge you expect to extract from your sequences. What are*

*you primarily interested in: sequencing, timing, duration, quantum?*

Sequencing, timing, duration (as much as I can !!!)

Could you please tell me what quantum is ?