# C. Interstellar First Contact (1/2) [10 points]

*It's hard enough to translate between languages when you understand both languages. It's harder still when you only understand one. But what do computers do? They don't truly understand either language. To illustrate the challenge that computers face, Kevin Knight posed this classic puzzle (Knight 1997): given two equivalent texts in two unknown alien languages, how would you go about translating one to another?*

It is the year 2354 AD. Our scientists have been eavesdropping on messages between two alien civilizations for a very long time, but we have never met either. The closest aliens, the Centauri, have finally begun to communicate with us. Their first message was a message of peace, "Farok crrok hihok yorok clok kantok ok-yurp."

Now, we know that the Centauri have been in contact for some time with the Arcturan race, who live in another solar system. We have never had contact with the Arcturans, but newly developed technology makes it possible for us to send them a message. We would like to send them, first, a message of peace, but because we do not understand their language, this is not an easy task.

Luckily, we have intercepted communications from the Centauri that include both languages. Here are 12 sentences in Centauri and their 12 translations in Arcturan. Unfortunately, because we have only been eavesdropping, their meaning is unknown. However, we do know that the sentence pairs on each line are translations of each other. We want to use this information to translate the original peace message from the Centauri and then send this to the Arcturans. Your assignment will be to do this translation.

| CENTAURI | ARCTURAN |
|---|---|
| ok-voon ororok sprok. | at-voon bichat dat. |
| ok-drubel ok-voon anok plok sprok. | at-drubel at-voon pippat rrat dat. |
| erok sprok izok hihok ghirok. | totat dat arrat vat hilat. |
| ok-voon anok drok brok jok. | at-voon krat pippat sat lat. |
| wiwok farok izok stok. | totat jjat quat cat. |
| lalok sprok izok jok stok. | wat dat krat quat cat. |
| lalok farok ororok lalok sprok izok enemok. | wat jjat bichat wat dat vat eneat. |
| lalok brok anok plok nok. | iat lat pippat rrat nnat. |
| wiwok nok izok kantok ok-yurp. | totat nnat quat oloat at-yurp. |
| lalok mok nok yorok ghirok clok. | wat nnat gat mat bat hilat. |
| lalok nok crrok hihok yorok zanzanok. | wat nnat arrat mat zanzanat. |
| lalok rarok nok izok hihok mok. | wat nnat forat arrat vat gat. |

n a c l o
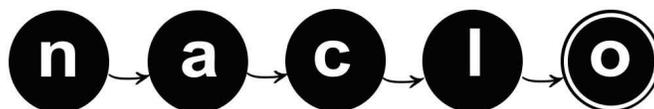
# C. Interstellar First Contact (2/2)

**C-1** Let's start with the first Centauri word: "farok". This word occurs in two of our Centauri sentences. Given that these sentences' Arcturan translations only have one word in common with each other, we can assume that this word is the translation for "farok". Which word it is?

farok

**C-2** Do the same thing for "hihok" and "yorok". For "yorok" you will need to make some assumptions about word ordering.

hihok

yorok

**C-3** The Centauri word "clok" only occurs once. However, you can figure out its Arcturan translation in another way.

clok

**C-4** Try to use the processes from the previous assignments to complete as much as possible of the following table.

crrok

kantok

ok-yurp

**C-5** Complete the translation of "farok crrok hihok yorok clok kantok ok-yurp." Keep in mind that Centauri and Arcturan sentences can have a different word order. There may be more than one correct reply.

**C-6** After some years a reply message is received in Arcturan. It reads, "Totat nnat forat arrat mat bat." Translate this message into Centauri. There may be more than one correct reply.

n a c l o