Problem 1

t	Packets sent by A	Packets sent by B
0	SYN, seq=20000, win=8000	
1		SYN,ACK, seq=5000, ack=20001, win=8000
2	ACK, seq=20001, ack=5001, win=8000	
3	ACK, seq=20001, ack=5001, win=8000, data=DataA1	
4		ACK, seq=5001, ack=20101, win=7900
5	ACK, seq=20101, ack=5001, win=8000, data=DataA2	
6		ACK, seq=5001, ack=20201, win=7800
7		[*B calls* `read()`, *which returns 200 bytes*]
8	ACK, seq=20201, ack=5001, win=8000 data=DataA3	
9		ACK, seq=5001, ack=20301, win=7900
10		ACK, seq=5001, ack=20301, win=7900, data=`DataB1`
11	ACK, seq=20301, ack=5101, win=7900	
12		FIN,ACK, seq=5101, ack=20301, win=7900
13	. . .	

Problem 2

(a) Six seconds elapsed between the two because the TTL decreased by 6 seconds for each response.

(b) The second query took less long because the result have been cached by a server that is closer to you than the authoritative DNS server.

(c) It's possible that the two answers are in different orders because one server is busier than the other and the returned order is used for load-balancing.

(d) You can use DNS resolvers that are responsible for a large number of users, so when someone in one region visits a site, that site address is cached and will be quicker for everyone else to look up.

(e) The ISP typically runs a DNS resolver, so they can modify or delete a query as it happens and prevent a user from accessing the site that they are trying to access.

(f) Google uses a shorter TTL because people search many different things, and some searches are unique. So, it doesn't make much sense to cache the addresses for very long, since the same search might not come up again