۴) الف) عمق، در مورد character encodings ۸

یکی از روش‌های عکس دادن دیجیتالی کاراکتر‌ها، وسیله الویا اعداد ست که هر کاراکتر نسبت داده می‌شود. درواقع درانیای نوعی قرارداد برای درست نوشته در تبادل اطلاعات است (مثلاً برای اینکه داده‌های ارسالی از طرف کامپیوتر X به Y قابل فهم و مشترک باشد) ماشین درانیای مورس، اسلی، یونیکد درقسمت ب توضیح داده شده‌اند.

ب) تفاوت تعداد کاراکتر‌های مجاز برای الفبای فارسی و انگلیسی است از جنبه‌ی و چه بستری باهم دارند؟ کاراکتر‌های انگلیسی براساس الویا درانیای به نام "ASCII" طراحی شده و ۸ بیت (۱ بایت) از حافظ را اسغال می‌شدند درحالت کاراکترع فارسی براساس درانیای UTF-8 که ۱۶ بیتی هستند (۲ بایت) طراحی شده‌اند پس هر کاراکتر فارسی دو برابر هر کاراکتر انگلیسی حجم اسغال می‌شند پس تعداد کاراکتر‌های مجاز برای الفبای فارسی دو برابر تعداد کاراکتر‌های مجاز انگلیسی است.

پ) مقایسه بین دو استاندارد کدگذاری "UTF" و "ASCII" :

ASCII (مخفف) (American Standard Code for Information Interchange
استاندارد کدگذاری آمریکایی برای تبادل اطلاعات است. کاراکترست اساسی و نوع
توسعه یافته می‌شود نوع ۷ بیتی که تعداد $۲^۷ = ۱۲۸$ کاراکتر مختلف است. (از صفر تا ۱۲۷)
نوع دیگر آن ۸ بیتی است (اساس توسعه یافته) و دارای $۲^۸ = ۲۵۶$ کاراکتر مختلف
است (از صفر تا ۵۵۶) حالت توسعه یافته از حالت استاندارد جداست. جایی از صفر تا ۱۲۷
کاراکتر اول درست مثل حالت استاندارد و فقط یک کاراکترهایی به شماره‌های از ۱۲۸ تا ۲۵۵
اضافه شده‌است. کاراکترهای اضافی در اول هیچ استانداردی نبرده و مبنی است در
تقسیم‌های مختلف فرق کنند. اما ۳۲ کاراکتر اول یعنی (۵ تا ۳۱) و آخرین کاراکتر (۱۲۷)
به‌عنوان کاراکترهای کنترلی شناخته می‌شوند که غیرقابل چاپی هستند و فقط برای کنترل
متن مورد استفاده قرار می‌گیرد. بقیه کاراکترها یعنی از ۳۲ تا ۱۲۶ قابل چاپی هستند.
مثلاً حالا علامت‌ها، حروف و اعداد انگلیسی هستند. کاراکترهای ۱۲۸ تا ۲۵۵ هم قابل
چاپی هستند.

UTF (مخفف) (unicode transformation format
UTF طور مخصوص UTF_8 یک uncoding با یک تا چهار بایتی است. (این
استاندارد دیگری برای کدگذاری می‌باشد. این کدگذاری جزء سمت‌های ۸ و ۱۶، ۳۲ بیتی دارد.
یکی از مزیت‌های این کدگذاری این است که سازگاری کامپیوترم ساده و پردازش کند دارد.
(مقدار مصرف حافظه CPU در مقابل مصرف بیشتر حافظه یا حجم ذخیره‌سازی. آن مثال)
UTF_16 برای هر کاراکتر چهار بایت دارد (هرچند کاراکتر بیشتر از ۱ بایت نیاز داشته باشد)
اکثر این کدگذاری در بعضی از سیستم‌ها و محیط‌های سیستم عامل‌ها استفاده می‌شود.